

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 05:52:06 ; Search time 1253.82 Seconds
(without alignments)

8607.626 Million cell updates/sec

Title: US-09-895-435A-4

Sequence: 1 gggggtacggccgacgacatcg.....gacggcgaggaatacatcatg 249

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 3470272 segs, 21671516935 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	249	100.0	249	6	AX671525	AX671525 Sequence
2	244	98.0	244	6	AX671522	AX671522 Sequence
3	241	96.8	4149	1	BTM4430	X07651 Bacillus th
4	241	96.8	9672	1	BTM42XX	X13481 Bacillus th
5	118	47.4	118	6	AX671523	AX671523 Sequence
6	114.4	45.9	116	6	AX671528	AX671528 Sequence
7	114.4	45.9	116	6	AX671530	AX671530 Sequence
8	64	25.7	116	6	AB064263	AB064263 plasmodu
9	53	21.3	2015	3	PFMA13P2	AL034558 plasmodu
10	53	21.3	158548	3	PFMA13P2	AC005506 plasmodu
11	53	21.3	205429	2	AC005506	AC005506 plasmodu
12	53	21.3	251132	3	AE014846	AE014846 plasmodu
13	52.8	21.2	609	8	PRU251061	AE014846 plasmodu
14	52	20.9	1018	8	AF402569	AF402569 Clavija s
15	51.8	20.8	29150	3	AC115683	AC115683 Dictyoste
16	51.8	20.8	83332	9	AC095051	AC095051 Homo sapi
17	51.8	20.8	343050	3	PPA929353	AE014843 plasmodu
18	51.4	20.6	271546	3	AE014843	AE014843 plasmodu
19	51.2	20.6	61869	2	AC104381	AC104381 Homo sapi
20	51	20.5	124	6	AX671531	AX671531 Sequence
21	51	20.5	60529	2	AC144825	AC144825 Dario rex
22	51	20.5	90859	8	AC005561	AC005561 Arabidops
23	51	20.5	252632	3	AE014818	AE014818 plasmodu
24	50.8	20.4	203958	5	AL954339	AL954339 Zebrafish
25	50.4	20.2	110000	2	PFMA13_22	Continuation (23 o
26	50.2	20.2	6550	6	AX344862	AX344862 Sequence
27	50.2	20.2	99858	9	HSJ105949	AL116524 Human DNA
28	50.2	20.2	103942	2	AC091859	AC091859 Homo sapi
29	50.2	20.2	110000	2	AC117408_1	Continuation (2 of
30	50.2	20.2	113372	9	AL161613	AL161613 Human DNA
31	50.2	20.2	148263	9	AC091988	AC091988 Homo sapi
32	50.2	20.2	172043	9	AC026797	AC026797 Homo sapi
33	50.2	20.2	173181	9	EX005037	EX005037 Human DNA
34	50.2	20.2	182366	9	AL591926	AL591926 Human DNA
35	50.2	20.2	188062	9	AL162731	AL162731 Human DNA
36	50.2	20.2	250029	3	AE014820	AE014820 plasmodu
37	50.2	20.2	250707	3	AE014848	AE014848 plasmodu
38	50.2	20.2	345552	3	PPA929354	AL929354 plasmodu
39	50	20.1	11308	1	AE007654	AE007654 Clostridi
40	50	20.1	163316	9	AL161420	AL161420 Human DNA
41	50	20.1	182083	2	BX539308	BX539308 Dario rex
42	49.8	20.0	32392	6	AX392737	AX392737 Sequence
43	49.8	20.0	50000	6	AX392733	AX392733 Sequence
44	49.8	20.0	232392	14	AF250284	AF250284 Amsacta m
45	49.8	20.0	232392	14	AF250284	AF250284 Amsacta m

ALIGNMENTS

RESULT 1	AX671525	249 bp	DNA	linear	PAT 27-MAR-2003
LOCUS	AX671525	Sequence 4 from Patent WO03004652.			
DEFINITION	AX671525				
ACCESSION	AX671525				
VERSION	AX671525.1	GI:29329875			
KEYWORDS					
SOURCE					
ORGANISM					
	Bacillus thuringiensis				
	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus				
	cereus group.				
REFERENCE	1				
AUTHORS	Stewart, F.A., Zhang, Y. and Hallet, B.				
TITLE	Use of a tyrosine recombinase for genetic engineering				
JOURNAL	Patent: WO 03004652-A 4 16-JUN-2003.				

Pred. No. is the number of results predicted by chance to have a

The European Molecular Biology Laboratory (DE) ; L'Universite
Catholique De Louvain (BE)
Location/Qualifiers
1..249
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"

ORIGIN

Query Match 100.0%; Score 249; DB 6; Length 249;

Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCGCATTTTCGAAAAAACCAGCTAGAAATCAGATTAAAAATCAG 60
DB 1 GGGGTACCGCGCATTTTCGAAAAAACCAGCTAGAAATCAGATTAAAAATCAG 60
QY 61 AAAATATCATTTATCTTGCACATACATCTTTTATACAAAAATATACAA 120
DB 61 AAAATATCATTTATCTTGCACATACATCTTTTATACAAAAATATACAA 120
QY 121 CACATATTAATTTGTTTATAGTGTATTAATTAATTAATCTAGGGTTTAA 180
DB 121 CACATATTAATTTGTTTATAGTGTATTAATTAATTAATCTAGGGTTTAA 180
QY 181 AACACATTTATGATTAATTAATTAATCTTTAGACGCAACATTTTATAG 240
DB 181 AACACATTTATGATTAATTAATTAATCTTTAGACGCAACATTTTATAG 240
QY 241 AATCAGATG 249
DB 241 AATCAGATG 249

RESULT 2

AX671522 244 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 1 from Patent WO03004652.
DEFINITION AX671522
ACCESSION AX671522
VERSION AX671522.1 GI:29329872

KEYWORDS

Bacillus thuringiensis
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.

REFERENCE 1
AUTHORS Stewart, F.A., Zhang, Y. and Hallet, B.
TITLE Use of a tyrosine recombinase for genetic engineering
JOURNAL Patent: WO 03004652-A 1 16-JAN-2003;
The European Molecular Biology Laboratory (DE) ; L'Universite
Catholique De Louvain (BE)
Location/Qualifiers
1..244
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"

FEATURES

source
ORIGIN

Query Match 98.0%; Score 244; DB 6; Length 244;

Best Local Similarity 100.0%; Pred. No. 8.3e-32;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGTACCGCGCATTTTCGAAAAAACCAGCTAGAAATCAGATTAAAAATCAGAA 62
DB 1 GGTACCGCGCATTTTCGAAAAAACCAGCTAGAAATCAGATTAAAAATCAGAA 60
QY 63 AATATATCATTTATCTTGCACATACATCTTTTATACAAAAATATACAA 122
DB 61 AATATATCATTTATCTTGCACATACATCTTTTATACAAAAATATACAA 120
QY 123 CAAATATTAATTTGTTTATAGTGTATTAATTAATTAATCTAGGGTTTAA 182
DB 121 CAAATATTAATTTGTTTATAGTGTATTAATTAATTAATCTAGGGTTTAA 180

QY 183 CACATTTATGATTAATTAATTAATCTTTAGACGCAACATTTATAGCGCGAGAAA 242
DB 181 CACATTTATGATTAATTAATTAATCTTTAGACGCAACATTTATAGCGCGAGAAA 240

ORIGIN

Query Match 243 TCAC 246

DB 241 TCAC 244

RESULT 3
BTNN4430 449 bp DNA linear BCT 12-SEP-1993
LOCUS BTNN4430
DEFINITION Bacillus thuringiensis transposon Th4430.
ACCESSION X07651
VERSION X07651.1 GI:40347
KEYWORDS plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 449)
Mahillon, J. and Lereclus, D.
Structural and functional analysis of Th4430: identification of an
integrase-like protein involved in the co-integrate-resolution
process
EMBO J. 7 (5), 1515-1526 (1988)
MEDLINE 88312602
PUBMED 2842151
Data kindly reviewed (03-APR-1999) by Lereclus D.
Location/Qualifiers
1..4149
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="H1.1"
/db_xref="taxon:1428"
/map="plasmid PG12"
/clone_lib="pBR322"
1..38 "inverted repeat A"
/note="inverted repeat A"
247..1101
/note="unnamed protein product; put. resolvase Tnp I (AA 1
-284)"
/codon_start=1
/transl_table=1
/protein_id="CA30491.1"
/db_xref="GI:40348"
/db_xref="GOL:P10020"
/db_xref="SWISS-PROT:P10020"
/translation="MDVAKQFSSYLKQENKENTVQGYTSIGIRQYIKMEGYSYDRKIT
KLYRONILLEYISLKNVMEINAKSINKISLAKFNFLLQKSOQOQVILKIDMITKY
QTVVASPQIVLIDVKKFELQVLEDDNKKRYALITLLAYGVISLAKDIFNIO
TGEICIRSGKSKORIVLINSKVSALKDYLIDKVTSTAHSEYPLISKREKLDRT
VYNRERKSYSVVITPHQIRHPCMAIEKXSHIEVANGSHNITTLTYTPNQDQ
LKNRKELL"
1120..4083
/note="unnamed protein product; transposase Tnp A (AA 1 -
987)"
/codon_start=1
/transl_table=1
/protein_id="CA30492.1"
/db_xref="GI:40349"
/db_xref="GOL:P10021"
/db_xref="SWISS-PROT:P10021"
/translation="MGVQQLSEAKQNELMDLSRLTWDLVTRTSKDLHLILKLR
RGYNLGLPALVLRYPGMSLLEYKQIPQVVAIVTSRLIPPEELVYAKGNTLV
EHLGEIRTEYQONFSEYKETLLQPLVOQAMNNNTLYLEITISLRKTKVILPAM
VYIEDIWEAKQADQKYSKILHDLQVOEQDQDALPLINKSPILAMIKQVAPQ
SPESFLKVIDLQFQKIGLIDITTKINTNRRLQALSGSEYVAPRRNENYVSM
LVSFLLEITDPLIDVLEIHLRLMMNIOKSKKQODIIOANGKXNKILLOETVQ
TLEAKETGKAPALADPVMWEMWMSVPAKOLSRPLAYVDYDILINTSVYVRSA
PLILSLHFRATKGEPLQALDITHELNETGKRVHSGAPLHVSVMKQHYVDDDS

REFERENCE 1
AUTHORS Mahillon, J. and Lereclus, D.
TITLE Structural and functional analysis of Th4430: identification of an
integrase-like protein involved in the co-integrate-resolution
process
EMBO J. 7 (5), 1515-1526 (1988)
MEDLINE 88312602
PUBMED 2842151
Data kindly reviewed (03-APR-1999) by Lereclus D.
Location/Qualifiers
1..4149
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="H1.1"
/db_xref="taxon:1428"
/map="plasmid PG12"
/clone_lib="pBR322"
1..38 "inverted repeat A"
/note="inverted repeat A"
247..1101
/note="unnamed protein product; put. resolvase Tnp I (AA 1
-284)"
/codon_start=1
/transl_table=1
/protein_id="CA30491.1"
/db_xref="GI:40348"
/db_xref="GOL:P10020"
/db_xref="SWISS-PROT:P10020"
/translation="MDVAKQFSSYLKQENKENTVQGYTSIGIRQYIKMEGYSYDRKIT
KLYRONILLEYISLKNVMEINAKSINKISLAKFNFLLQKSOQOQVILKIDMITKY
QTVVASPQIVLIDVKKFELQVLEDDNKKRYALITLLAYGVISLAKDIFNIO
TGEICIRSGKSKORIVLINSKVSALKDYLIDKVTSTAHSEYPLISKREKLDRT
VYNRERKSYSVVITPHQIRHPCMAIEKXSHIEVANGSHNITTLTYTPNQDQ
LKNRKELL"
1120..4083
/note="unnamed protein product; transposase Tnp A (AA 1 -
987)"
/codon_start=1
/transl_table=1
/protein_id="CA30492.1"
/db_xref="GI:40349"
/db_xref="GOL:P10021"
/db_xref="SWISS-PROT:P10021"
/translation="MGVQQLSEAKQNELMDLSRLTWDLVTRTSKDLHLILKLR
RGYNLGLPALVLRYPGMSLLEYKQIPQVVAIVTSRLIPPEELVYAKGNTLV
EHLGEIRTEYQONFSEYKETLLQPLVOQAMNNNTLYLEITISLRKTKVILPAM
VYIEDIWEAKQADQKYSKILHDLQVOEQDQDALPLINKSPILAMIKQVAPQ
SPESFLKVIDLQFQKIGLIDITTKINTNRRLQALSGSEYVAPRRNENYVSM
LVSFLLEITDPLIDVLEIHLRLMMNIOKSKKQODIIOANGKXNKILLOETVQ
TLEAKETGKAPALADPVMWEMWMSVPAKOLSRPLAYVDYDILINTSVYVRSA
PLILSLHFRATKGEPLQALDITHELNETGKRVHSGAPLHVSVMKQHYVDDDS

repeat_unit

CDS

CDS

```

NINRYVELALTELNRHRSRSDIFVSGSRHKAADYLLPYDEMNEVSNIPNGLTAP
LKADYITDRINRLNHEMLSKNSKLEGVDSOGKLHVERLDKGTPEAKAFSKLL
HSMLEPRIKLTDLLIEVASWTFHDFQFIHASTNOSPDEONIVLATMAMGNIIGLR
MAEATPGISYROMANASOMRWDDAMRAOSILVNPKEOKLSSYNGDSTSSDPMR
LSIAVRSLSHADSNSHYTGKGGTIRYVSDDLSAHHVYVITTNARDLHYLDLHHE
TDLKIEHYITDAGTDOVFALTHLGRFAPRIRLADITKLSIPGEEYENVOALL
KGINVKLIKENYEDIRRLAYSVGTGKVSALLIGKLSYARQNKLATLGMGRLEK
TLFTLDYISNKAVERRVQKLNKEAINALARTIFGQGEFERRALODOLQASALN
IINAIISVNTVMEKAVEELKARGEFREDIMPYAMPJGWEHINFLGEYFEGJLHDTG
OMNLRPLRIKEPFYS"
repeat_unit
4112..4149
/note="inverted repeat A'"

ORIGIN
Query Match          96.8%; Score 241; DB 1; Length 4149;
Best Local Similarity 98.0%; Pred. No. 1.4e-31;
Matches 244; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGGTACCCGACGCTTCGGAAAAAACCCGCTAAGAAATCAAGTTTAAATCAG 60
DB 1 GGGGTACCCGACGCTTCGGAAAAAACCCGCTAAGAAATCAAGTTTAAATCAG 60
QY 61 AAATATATCATTTATCTTGACATCATGTTCTTTTATACAAAAATATATCA 120
DB 61 AAATATATCATTTATCTTGACATCATGTTCTTTTATACAAAAATATATCA 120
QY 121 CACAAATTTAATGTGTGTATAGGTGTATTAATTAATTAATCTTAAAGGGTTTAA 180
DB 121 CACAAATTTAATGTGTGTATAGGTGTATTAATTAATTAATCTTAAAGGGTTTAA 180
QY 121 CACATATTTAATGTGTGTATAGGTGTATTAATTAATTAATCTTAAAGGGTTTAA 180
DB 121 CACATATTTAATGTGTGTATAGGTGTATTAATTAATTAATCTTAAAGGGTTTAA 180
QY 181 AACCAATTTTCGATTAATTAATCTTTTACAGCACAACATTTTACGCGGAGGA 240
DB 181 AACCAATTTTCGATTAATTAATCTTTTACAGCACAACATTTTACGCGGAGGA 240
QY 241 AATCACATG 249
DB 241 AATCACATG 249

RESULT 4
BTPG12XX          9672 bp DNA linear BCT 07-JUL-2002
LOCUS              Bacillus thuringiensis plasmid pg12 with transposon Tn4430.
DEFINITION          X13481
ACCESSION            X13481
VERSION              X13481.1 GI:3171732
KEYWORDS              plasmid, pg12; recombinase; resolvase; transposase;
                      transposon; unidentified reading frame.
SOURCE
ORGANISM              Bacillus thuringiensis
                      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
                      cereus group.
REFERENCE
AUTHORS              1 (bases 1 to 6999)
                      Mahillon, J. and Seurinck, J.
TITLE                Complete nucleotide sequence of pg12, a Bacillus thuringiensis
                      plasmid containing Tn4430
JOURNAL              Nucleic Acids Res. 16 (24), 11827-11828 (1988)
MEDLINE              89098342
PUBMED              3211758
AUTHORS              2
                      Mahillon, J.
TITLE                Direct Submission
JOURNAL              Submitted (04-NOV-1988) Mahillon J., Plant Genetics Systems, J
                      Plateaustraat 22, B-9000 Gent, Belgium
REMARK              revised by [3]
AUTHORS              3 (bases 1 to 9672)
                      Hoflack, L.
TITLE                Direct Submission
JOURNAL              Submitted (24-MAR-1998) Hoflack L., Plant Genetics Systems, J
                      Plateaustraat 22, B-9000 Gent, Belgium
COMMENT              On Jun 2, 1998 this sequence version replaced gi:40316.
FEATURES
source              1..9672
                      /organism="Bacillus thuringiensis"

repeat_unit
repeat_region
repeat_region
repeat_unit
CDS
/mol_type="genomic DNA"
/strain="thuringiensis H.1."
/db_xref="taxon:11428"
/plasmid="pg12"
493..4651
/transposon="Tn4430"
493..497
/note="direct repeat 1"
498..535
/note="inverted repeat A"
744..1598
/codon_start=1
/transl_table=11
/product="TnpI resolvase"
/protein_id="CAA31832.1"
/db_xref="GI:40317"
/db_xref="GOA:P10020"
/db_xref="SWISS-PROT:P10020"
/translation="MDYAKQSSSYLKKENKENTVQGYTSGRQYIKPFESYDRLK
KLVRNLIIEYISYLNKVRMLNKSINHKISLAKNEFLIQSGQDDVILKTDIKV
QTVVASPTQIVELDYKFLQSVLEEDNNRYALATLAVTGVRSLEALS IKNDENLQ
TGEICIRSGKGGKQRIVLNSKVSIAIDYLDIKTVEFASHPYLSFKKEKLDRT
VYNRIFKSYSNVITPHQJLHFQCTNAIEKRSIHEVANAQGHSHITLLVTNPNQLO
LKRMELL"
1617..4580
/codon_start=1
/transl_table=11
/product="TnpA transposase"
/protein_id="CAA31833.1"
/db_xref="GI:40318"
/db_xref="GOA:P10021"
/db_xref="SWISS-PROT:P10021"
/translation="WGVKQLSEAOGRNLMDSRLTENDVTFPTFSKXDDLILKHR
RGYNRLGFLRLVILIRPQMSLTXEKDIPQYVAVAVTSRLRIPPEFLVAKRGVTLW
EHJGIRTEYGVONFSSEKXETLQFLVQAOQDNNNTLEITISLTKVILIPAM
VYIEDIWEAKQAOQKXYSILHDDVQAOQDNLALPLTINGKSPALMKVILIAQ
SPESFLAKVIDIRLPYQKGLTIDTKIKTNRLRLARLGSKTEPYAPRPEVKAQSM
LVSFLIETNODLIDYAIERHIDLMNMLQTKGKEDEDIQOANGKLNKILIOFIVCG
TILBAKETKQDAFALDVEYMSNMWVESYEAKQSRPLANDYDLINTRSYVARYA
PTLRSLSHFRATSGEPVQALDITHEINERTKRPVPGAPLHVFNSRQKHYDDG
NINRYVELALTELNRHRSRSDIFVSGSRHKAADYLLPYDEMNEVSNIPNGLTAP
LKADYITDRINRLNHEMLSKNSKLEGVDSOGKLHVERLDKGTPEAKAFSKLL
HSMLEPRIKLTDLLIEVASWTFHDFQFIHASTNOSPDEONIVLATMAMGNIIGLR
MAEATPGISYROMANASOMRWDDAMRAOSILVNPKEOKLSSYNGDSTSSDPMR
LSIAVRSLSHADSNSHYTGKGGTIRYVSDDLSAHHVYVITTNARDLHYLDLHHE
TDLKIEHYITDAGTDOVFALTHLGRFAPRIRLADITKLSIPGEEYENVOALL
KGINVKLIKENYEDIRRLAYSVGTGKVSALLIGKLSYARQNKLATLGMGRLEK
TLFTLDYISNKAVERRVQKLNKEAINALARTIFGQGEFERRALODOLQASALN
IINAIISVNTVMEKAVEELKARGEFREDIMPYAMPJGWEHINFLGEYFEGJLHDTG
OMNLRPLRIKEPFYS"
4609..4646
/note="inverted repeat A'"
4647..4651
/note="direct repeat 1"
4859..5473
/note="unnamed protein product; ORF 1"
/codon_start=1
/transl_table=11
/protein_id="CAA31836.1"
/db_xref="GI:3171733"
/db_xref="GOA:P10022"
/db_xref="SWISS-PROT:P10022"
/translation="MTNITDFGHVKNARIRDFENKEKFKODHNGINGOEVNAQMV
LSKATGKEIFIGTKRSPQSKFAQFIQDDNDVYALNENAFDEEMLFLIRLOEIOF
KNCIYNDHSHNALPMSQKOIADLKQKLSIRIVNSLIVNSLVQKGYIVANGKPPGVK
ARYVALFINPNTIYGERNVTETIKALPMNSKSLFKPEPILAF"
6037..6849
/note="unnamed protein product; ORF 2"
/codon_start=1
/transl_table=11
/protein_id="CAA31837.1"
/db_xref="GI:3171734"
/db_xref="GOA:P10023"

```

/db_xref="SWISS-PROT:P10023"
 /translation="MKLLISFIALFFICGPNLKAAEEIIDYOSLYNQAOIEGVL
 DQNSVYENMLKONKEEMPIYODGLKQGLLEPLSENMKLNNTGQATGIDLEFD
 DVTGSGMGFTLKAGDIFITNATSSAGIYGHAIAANGDYIILHMGAGGNOGLSTS
 NMOKYTAGSGWTYVAILKQOTLARADYARADYRPFYTSATIKNYIYLDGITHYQ
 KNPTYSKLYVQALYFSGSRNVQAVSIVTPIGLDITFSAPRSVAKTY"
 6896..7240
 /note="unnamed protein product; ORF 3"
 /codon_start=1
 /transl_table=1
 /protein_id="CAJ1834.1"
 /db_xref="GI:40321"
 /db_xref="GOA:P10024"
 /db_xref="SWISS-PROT:P10024"
 /translation="MMAIFISIVLVSPVLEFNRAQAKSIYNYIARQIGESQLYID
 FHKDFKGGWLVAVYEGENDIYVEYSYDCKVNFQAYNSSEKAIKKKMGSGGLTE
 IEMKKIKYPLQ"
 7829..9166
 /codon_start=1
 /transl_table=1
 /product="put. G12 site-specific recombinase"
 /protein_id="CAJ1835.1"
 /db_xref="GI:40322"
 /db_xref="GOA:P10025"
 /db_xref="SWISS-PROT:P10025"
 /translation="MMKFAIHHMKRGISDVQGIQKNGRQGSKNLDIDYKSN
 YDLNQKIRYESTKQISERVAKRKRAVSVLSEFVYASPDYMSLSLEQKRYF
 ESSLDIFQKRYGKQNTYAMVHMDAPPHAIQVMPITENRLSADQMFRLKLSIQ
 QDPLEMEKGFVDKRGESGKHLSPQAREKQDLEVEQLSNVTHIKTVFETH
 NLOQNTNYIEKQNETLQKIQOQPLSLDKIKKEKQEFEPFQNOIDPKPYMSYLRB
 TKTEVTKLFKPEITEKGTGNIIVTSEORNDMEKNAIVIKYERLQKIDLVKE
 NOSLRDNKYLEETIKNNLAKHSYKQNELEVENKELTEIGTCAHTRDQNMVK
 VLYQTKVPEQKFAFRGLIKNELDKGVNDQPEPHEHTIEISRQKIDMER"

ORIGIN

Query Match 96.8%; Score 241; DB 1; Length 9672;
 Best Local Similarity 98.0%; Pred. No. 1.2e-31;
 Matches 244; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTCCGAAAAAACCCAGCTAGAGAAATAGAGTTAAAAATCAG 60
 DB 498 GGGGTACCGCCAGCATTTCCGAAAAAACCCAGCTAGAGAAATAGAGTTAAAAATCAG 557
 QY 61 AAAATATATCATTTATTCCTTGACACATCATGTCCTTTTATATACAAAAATATATCA 120
 DB 558 AAAATATATCATTTATTCCTTGACACATCATGTCCTTTTATATACAAAAATATATCA 617
 QY 121 CACATATTAATGCTGTGATAGAGTGTATTAATTAATTAATCTAGGGGTTTAAGC 180
 DB 618 CACATATTAATGCTGTGATAGAGTGTATTAATTAATTAATCTAGGGGTTTAAGC 677
 QY 181 AACCAATTTATCGATAAATAATCTTTAGACGCAACACAAATTTATAGCGCGAGGA 240
 DB 678 AACCAATTTATCGATAAATAATCTTTAGACGCAACACAAATTTATAGCGCGAGGA 737
 QY 241 AATCACATG 249
 DB 728 AATCACATG 746

RESULT 5

LOCUS AX671523 118 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 2 from Patent WO03004652.
 ACCESSION AX671523
 VERSION AX671523.1 GI:29329873
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus thuringiensis
 Bacillus thuringiensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.
 REFERENCE 1
 AUTHORS Stewart, F. A., Zhang, Y. and Hallet, B.

TITLE Use of a tyrosine recombinase for genetic engineering
 JOURNAL Patent: WO 03004652-A 2 16-JAN-2003;
 The European Molecular Biology Laboratory (DE) ; L'Universite
 Catholique De Louvain (BE)
 FEATURES
 source location/Qualifiers
 1..118
 /organism="Bacillus thuringiensis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1428"

ORIGIN

Query Match 47.4%; Score 118; DB 6; Length 118;
 Best Local Similarity 100.0%; Pred. No. 1.4e-10;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 TAAATCAACAAATTAATGCTGTGATTAAGGTATTAATTAATTAATTAATCTAGCGG 172
 DB 1 TAAATCAACAAATTAATGCTGTGATTAAGGTATTAATTAATTAATTAATCTAGCGG 60
 QY 173 TTTAAGCAACAAATTTATCGATTAATTAATTAATCTTTTAAAGCAACAAATTTATAG 230
 DB 61 TTTAAGCAACAAATTTATCGATTAATTAATTAATCTTTTAAAGCAACAAATTTATAG 118

RESULT 6

LOCUS AX671528 116 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 7 from Patent WO03004652.
 ACCESSION AX671528
 VERSION AX671528.1 GI:29329878
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus thuringiensis
 Bacillus thuringiensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.

REFERENCE 1
 AUTHORS Stewart, F. A., Zhang, Y. and Hallet, B.
 TITLE Use of a tyrosine recombinase for genetic engineering
 JOURNAL Patent: WO 03004652-A 7 16-JAN-2003;
 The European Molecular Biology Laboratory (DE) ; L'Universite
 Catholique De Louvain (BE)
 FEATURES
 source location/Qualifiers
 1..116
 /organism="Bacillus thuringiensis"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1428"

ORIGIN

Query Match 45.9%; Score 114.4; DB 6; Length 116;
 Best Local Similarity 99.1%; Pred. No. 5.8e-10;
 Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 113 TAAATCAACAAATTAATGCTGTGATTAAGGTATTAATTAATTAATTAATCTAGCGG 172
 DB 1 TAAATCAACAAATTAATGCTGTGATTAAGGTATTAATTAATTAATTAATCTAGCGG 60
 QY 173 TTTAAGCAACAAATTTATCGATTAATTAATTAATCTTTTAAAGCAACAAATTTAT 228
 DB 61 TTTAAGCAACAAATTTATCGATTAATTAATTAATCTTTTAAAGCAACAAATTTAT 116

RESULT 7

LOCUS AX671529 116 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 8 from Patent WO03004652.
 ACCESSION AX671529
 VERSION AX671529.1 GI:29329879
 KEYWORDS
 SOURCE
 ORGANISM
 Bacillus thuringiensis
 Bacillus thuringiensis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.
 REFERENCE 1

AUTHORS					
Stewart,F.A., Zhang,Y. and Hallet,B.					
TITLE					
Use of a tyrosine recombinase for genetic engineering					
JOURNAL					
Patent: WO 03004652-A 8 16-JAN-2003;					
The European Molecular Biology Laboratory (DE) ; L'Universite Catholique De Louvain (BE)					
FEATURES					
source					
1..116					
/organism="Bacillus thuringiensis"					
/mol_type="unassigned DNA"					
/db_xref="taxon:1428"					
ORIGIN					
Query Match					
Best Local Similarity 99.1%; Score 114.4; DB 6; Length 116;					
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
OY	113	TAATACACAACAATTATTTGCTGTGATTAGCGTATATAAATTAATCTAGGGG	172		
Db	1	TAATACACAACAATTATTTGCTGTGATTAGCGTATATAAATTAATCTAGGGG	60		
OY	173	TTTAACGCACAACTTTATCGATTAATAACTTTAGACGCACAACTTAT	228		
Db	61	TTTAACGCACAACTTTATCGATTAATAAATCCTTTAGACGCACAACTTAT	116		
RESULT 8					
AX671530 116 bp DNA linear PAT 27-MAR-2003					
LOCUS					
Sequence 9 from Patent WO03004652.					
DEFINITION					
AX671530					
ACCESSION					
AX671530.1 GI:29329880					
VERSION					
AX671530.1 GI:29329880					
KEYWORDS					
Bacillus thuringiensis					
Bacillus thuringiensis					
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.					
SOURCE					
Bacillus thuringiensis					
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.					
ORGANISM					
Bacillus thuringiensis					
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.					
REFERENCE					
1 Stewart,F.A., Zhang,Y. and Hallet,B.					
Use of a tyrosine recombinase for genetic engineering					
Patent: WO 03004652-A 9 16-JAN-2003;					
The European Molecular Biology Laboratory (DE) ; L'Universite Catholique De Louvain (BE)					
AUTHORS					
Stewart,F.A., Zhang,Y. and Hallet,B.					
TITLE					
Use of a tyrosine recombinase for genetic engineering					
JOURNAL					
Patent: WO 03004652-A 9 16-JAN-2003;					
The European Molecular Biology Laboratory (DE) ; L'Universite Catholique De Louvain (BE)					
FEATURES					
Location/Qualifiers					
1..116					
/organism="Bacillus thuringiensis"					
/mol_type="unassigned DNA"					
/db_xref="taxon:1428"					
ORIGIN					
Query Match					
Best Local Similarity 55.2%; Score 64; DB 6; Length 116;					
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;					
OY	113	TAATACACAACAATTATTTGCTGTGATTAGCGTATATAAATTAATCTAGGGG	172		
Db	1	TAATACACAACAATTATTTGCTGTGATTAGCGTATATAAATTAATCTAGGGG	60		
OY	173	TTTAACGCACAACTTTATCGATTAATAACTTTTAGACGCACAACTTAT	228		
Db	61	NNNNAACGCAACAATTTATNNNNNNNNNNNNNNNNNNNNNNNACGAACAATTTAT	116		
RESULT 9					
AB064263 2015 bp mRNA linear INV 26-DEC-2001					
LOCUS					
Plasmodium fulciparum fulm14 mRNA for glutaredoxin, complete cds.					
DEFINITION					
AB064263					
ACCESSION					
AB064263.1 GI:18143612					
VERSION					
AB064263.1 GI:18143612					
KEYWORDS					
Plasmodium fulciparum 3D7					
plasmodium fulciparum 3D7					
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
SOURCE					
Plasmodium fulciparum 3D7					
plasmodium fulciparum 3D7					
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
ORGANISM					
Plasmodium fulciparum 3D7					
plasmodium fulciparum 3D7					
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
REFERENCE					
1 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					

FEATURES	source
1. 2015	Location/Qualifiers
/organism="Plasmodium falciparum 3D7"	
/mol_type="mRNA"	
/strain="3D7"	
/db_xref="taxon:36329"	
/chromosome="3"	
/clone="XPFn269"	
/clone_lib="Plasmodium falciparum full-length cDNA library"	
/note="Identical sequence is also seen in clone XPFn675-erythrocytic-see also Acc#AB054090"	
1. 2015	
/gene="fmlm4"	
/155..1490	
/gene="fmlm4"	
/codon_start=1	
/product="glutaredoxin"	
/protein_id="BAB79691.1"	
/db_xref="GI:18143613"	
/translation="MAGTSRYKVKQWNIIEENIIVPAKTECPYCIKSIILKGYNLNSHHEHTEKPDMANIQYIKELTGKSSVPRIPINKDVYGGCDLVKENDSKLKERLKDLGVN"	
ORIGIN	
Query Match	21.3%; Score 53; DB 3; Length 2015;
Best Local Similarity	52.5%; Pred. No. 6.3;
Matches 116; Conservative	0; Mismatches 105; Indels 0; Gaps 0;
QY	22 AAAAAAAAAACGCGTAGAATAATGAGATTAAAAATCAGAAAATATATCATTTCTCTTG 81
Db	975 AAAAAAAAAAAAAAAAAAATTAATTAATTAATTAATTTGATTTATTTATTTAGTTTAAATA 1034
QY	82 ACACATCAGTGCTCTTTTATTAACAATAATATATCAACAAATATTAATGGTGTA 141
Db	1035 AAAAAATATATATCCCTCAATATTAATAAAATTTTAAAGAAATATATATTAATTCATT 1094
QY	142 TTAGGCTTAAATTAATTAATTAATTAATCTAGGGGTTTAAAGCAACAATTTATGATAATA 201
Db	1095 ATATATATATATATATATATATATATATATATATTAATTAATTAATTAATTAATTAATA 1154
QY	202 AATACTTTTAAAGCAACAACAATTTATAGCGCGGAGGAAA 242
Db	1155 ATGGCTGTACAGTGAAGCAGTTAAAAATGGGTAAACA 1195
RESULT 10	
PFML3P2/c	158548 bp DNA linear INV 29-JAN-2003
LOCUS	PFML3P2
DEFINITION	Plasmodium falciparum MAJSP2, complete sequence.
ACCESSION	AL003458 AL008982 AL008989 AL009007 AL009008 AL009009 AL009014
VERSION	AL010140 AL010150 AL010152 AL010156 AL010158 AL010211 AL022217
KEYWORDS	AL844502 Z97349 Z97350 Z98546
SOURCE	AL034558.3 GI:15375364
ORGANISM	Plasmodium falciparum 3D7
	Plasmodium falciparum 3D7
	Exkaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE
AUTHORS
1 (bases 1 to 158548)
Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Williams, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jaasal, B., James, S., McLean, J., Mole, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrett, B.G.
The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*
NATURE 400 (6744), 532-538 (1999)

TITLE
JOURNAL
MEDLINE
PUBMED
99376085
10448855

REFERENCE
AUTHORS
2
Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, C., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, S., Larke, N., Kechrou, A., Knights, A., Konfortov, B., Kyes, K.D., Johnson, D., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Mole, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, K., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrett, B.G.
Sequence of *Plasmodium falciparum* chromosomes 1, 3-9 and 13
NATURE 419 (6906), 527-531 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
12368867
3 (bases 1 to 158548)
Lawson, D., Bowman, S. and Barrett, B.
Direct Submission
Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UKXX

COMMENT
On or before Aug 30, 2001 this sequence version replaced
g1:2982545, g1:2982529, g1:2982547, g1:2982546, g1:2982548,
g1:2624372, g1:2982553, g1:2982560, g1:2894487, g1:2665315,
g1:2982565, g1:2982575, g1:2982578, g1:2982517, g1:2982518,
g1:2982519, g1:4493878.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.

FEATURES
Source
1. 158548
Location/Qualifiers
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="x3"
/clone="MAL3P2"
join(1030, 2237, 2429, 2502, 2645, 2685, 2826, 2941, 3046, 3173, 3308, 3388, 3508, 3571, 3699, 3737)
/gene="MAL3P2.1"
/note="synonym: PFC0160w"
join(1030, 2237, 2429, 2502, 2645, 2685, 2826, 2941, 3046, 3173, 3308, 3388, 3508, 3571, 3699, 3737)
/gene="MAL3P2.1"
/note="Revised: revised to 8 exon structure, revised: donor for exon 7
4 weak Pfam matches to entry PFC0023 ank, Ank repeat, Pfam match to entry PFC0159 zf-DHHC, DHHC zinc finger domain, score 66.70, E-value 4.8e-16
Similar to Plasmodium vivax pvlh14090 p pvlh14090w
SMALL:AA99466 (EMBL:AY003872) (611 aa) fasta scores: E(): 5.6e-172, 71.94% id in 613 aa, and to Mus musculus 503340611a1r1k protein 503340611a1r1k SMALL:Q9D3Q8 (EMBL:AK017158) (368 aa) fasta scores: E(): 1.7e-09,

gene
CDS
24.87% id in 197 aa, and to Rattus norvegicus small rec
srec SMALL:Q9YKRS (EMBL:AF228917) (366 aa) fasta scores:
E(): 8.4e-09, 27.95% id in 161 aa"
/codon_start=1
/product="binding protein, putative"
/protein_id="CA83988.3"
/db_xref="gi:23477021"
/db_xref="sp:EMBL:097224"
/translation="MNDSLSLSEVLEKQYELIKYAYODPIRLQIILIPYLNNDIE
MNSINILHMACYCGFTLEVKLISFNPDEKEDLVNDDPIYVAINSYSEVYLIL
KFGISILHFNKRRRMSPLTALICEFNDEKLELHLELLYNGVSENEHGTQRA
LELVKKNISITLQWLLTKENINHDYPGYGVTHLIVRHDDILRLDLCYGLNLY
YSSILEKNNVNPOLCTKREYELVYIILKKVYLNQKICSLKICKITTYAFEPFAL
NLIVYFNASHFSITNKHRSKLWITWTFPOQFLCMLYKFSKPFKMMFKMKK
NIQSKTKNSKPNPTNGFTKNAEYQLNNREBELQINKSLSTNFTLPIPHDQALS
NKYNDILINLEYQLSLYSQVQERINSLEBDYANLIVNPNICVTCNLTIPFV
HICAECEFCIVADHOCWVNDICIGIKQRFVWFICIFPLLYNYYVLYLPHQ
ATINAFESLIVILCNPIVNLFAFVLFARNTLITLVTFYHYKPKPHITDKYNT
DLRCMDPQNLMLKGLIKNIYFMSLNDPEYIRQCTMSDVCTLIISNT"
5341..12021
/gene="MAL3P2.2"
/note="synonym: PFC0165w"
5341..12021
/gene="MAL3P2.2"
/note="Revised: original entry split into PFC0165w and
PFC0166w
Similar to Plasmodium vivax pvlh14095 p pvlh14095w
SMALL:Q962J5 (EMBL:AT003872) blast scores: E(): 0.0,
score: 2249 43% id, and to Plasmodium falciparum
chloroexin-like redox-active protein fied SMALL:Q9NC62
(EMBL:AF234633) (179 aa) fasta scores: E(): 7.5e-47,
99.42% id in 175 aa"
/codon_start=1
/product="hypothetical protein, conserved"
/protein_id="CA83988.3"
/db_xref="gi:23477022"
/db_xref="sp:EMBL:097225"
/translation="MMNRKLIKLEKKNKNRNIKNVKEPDCHEMNNYSOCVYVK
KVENNDPMEKENTYNDNIIKRSQVEEEKKVIENYKESLSEPTFYRYSSNRK
IFLIINSNINLGNSCVNNDLIYIHNKKYEMANIPEKTKKKKKKKKKKKKKKK
KRLTNNKNTGPGSVTSMNDEKDYKYNQVDDDDDDDDDDDDDDDDDDDDDD
YEDDDDDDDDDDDNNKSYNSYDQNTKANKKKNPNNTSTIFTSNMFVCKE
RRKKKKKNSIDKRNITYSDDSDNNYETFTHSIDLPFLYSNBIKSNMLYQ
NEQFINDVDYFRPKFKNLFWNMDSCNSVTFKSLILNEYITRQKNDLVINYO
LYDVNNMYHRYNCTYRLYINDPDLINNKYKELIGQVYASHNTNNNNNDI
HNVNNNNMYNRREHLEKLEKKKKNYNNNNNNNDLSTTGDMLSNCKEYNE
SKYKLVIKQMTLKLHFNQDNTINLPSFNACTSNYKLDKYSYLLINNYLNK
SEHLINSRKKNYTLDHNNNTNNNNNTITKQVSNNNNNNNNDLSDLYSEFN
YLTKEIYENSTYSGLYKECLLIDILYILNTYGNLITKEYYKQHYEYNTYISN
IVQKKKRRKKKSKYKLYHSHKKNRQKQSTISFSQKATYNTMMNLHABE
IIPDQSMSSNISSSNMSSSKSHSATSIDIPMRNRKNNTYINRKOJNDLTY
FPPKXKLIPHTLNLVNEQNNNSIEIARELIVLQYIRKIQKLEIKRQSCNFTD
IFPCKLSINISAHILNIRLEENRNIYFPIKKNKSNFYNNNNNNNNNDNNFI
HNDLSLCLPDEYGLSTCTDPEEKKIFESINLPHKNIQDNLFPENIIEKAEISLY
KITSVLPLOIGKLINKLTIKKKKNINSKLLDILVILQELYNLDDINKKVLVY
LLKNTTLPFPKTYIYFGVYDTYKPEFPHKHTPTPYKONKHYENMYKVIK
KYIDSLNFWGAKYIILNDKVPFLKLSLNTITGKTIYDLSKSLINIQKTSYK
LIDQNEYNIENEDINDNTQNNRNSFLPQVDYKRIYVENINNOQYNSDHRNP
YFYNDLYKSERQTOHGMKQILINRKYKYNVNLVYDEYVEYHNDRKNSCFILYN
DKVYENLIHQHLYASKMFELYLKIDIEKIKYLFQIDSLYKLYLISHE
VLEKLYHKKNLILNKLPDIFSRSVYLHLYKKQYVAYLIDFISIMONVLLID
LRKQDSNEOGLNKKANNYEEDDDDDDDDDDDYNNNGENYNEEYDDEERNY
LNKNDYDDEQADDEEDDDDDDDDDDDDDDEYKYEYENDNDNDYDSDSHNNYNT
NENHDDQVHNSFRITKNPLMNDTNQKQYNEEYKTKERITIKNAELKKNKA
IKKKKKITKKKKNVSKNELERNKKNNTNNNNNNNNNNNNNNVDMGSPKLNK
NFIILKLNHNNSNNNNAVEYKQILISYKMPYNNIIPNNITPKYTLFVYLVK
NYEFLKSLVNNHNFVKNININECTNNMLCITRSCMIFELCYVLLQYLVNICE
YNNNNTKQKERTLPDDIHHIRYTLNNILKSPFMDNNIILCIKLISISIFPHLL
KPNEDKQCEIINTNGHGVSNVNLPTTHNNHNGYNNNNNNNNNNNNKHYKILQE
LLRQKAYISMTQNTIKATDFKAPFLITFYVANNIDYAHNELILQYDNYTMYK
KLSIQNPSFNNDINSEYVNDNDLDDSDTRQVKNINNGVNTNNNSKYTHKOT
SNNPTESSVQIKRKNLLANDNRNNTDILNNTKNDNNNNNNNNNGEYNTYREN
ITSEIKHNSPREYRNNIITNHTINNNSNHYQINTYSLHRSNNNNNNNNNNKQVSEL

HINDKSNFNKQNYNINDNNNNHYTNNNTSPKSNKRYNNHPIWENDYKX
 NINNNNNKOTILISENPKKXNHIPOKNNKDTSPHINNPNODENNNISPILNYSKX
 KENTSYNNNNKININLNIDSDISNRSKXNA"
 12790..13329
 /gene="PFC0165w"
 /note="synonym: fred"
 12790..13329
 /gene="PFC0165w"
 /note="Revised: split original spliced gene PFC0165w into
 PFC0165w and PFC0166w, fasta; plasmodium falciparum;
 thioredoxin-like redox-active protein; fred; length 179
 aa; id=100%; ungapped id=100%; E(=2.1e-69); 179 aa
 overlap; query 1-179 aa; subject 1-179 aa; fasta;
 TR:AA99466 (EMBL:AY003672); Plasmodium vivax;
 pvlh14100_p; pvlh14100w; length 179 aa; id=65.92%;
 ungapped id=67.42%; E(=1.5e-46); 179 aa overlap; query
 1-177 aa; subject 1-177 aa"
 /codon_start=1
 /product="thioredoxin-like redox-active protein, putative"
 /protein_id="QAD49085.1"
 /db_xref="GI:23477023"
 /db_xref="GOA:Q81224"
 /db_xref="SPTRMBL:Q81224"
 /translation="MACVDNPPKTYPNPDYAEYKRYANMYLYYQNNELKIDS
 YFKDYLGLFGASGCKYCVTFIDSLNIFKKNPVEIIVPDRYQYQSLKXNN
 FYALPDNYLYICKKQIKMLPSFMLITPNNNLYVQAAQLITDEYINNKSLINX
 IIRKTFQNNRPDLPFRN"
 complement(13855..15201)
 /gene="WAL3P2.3"
 /note="synonym: PFC0170c"
 complement(13855..15201)
 /gene="WAL3P2.3"
 /EC_number="2.3.1.-"
 /note="Similarity to lipamide acyltransferases. eg
 C.elegans lipamide acyltransferase (WP:ZK659.4) Blast
 Score: 582, sum P(1) = 9.2e-57; 33% identity in 430 aa
 overlap; similarity to P. vivax pvlh14105c, 59.908%
 identity (61.61% ungapped) in 434 aa overlap
 Pfam match to entry PFC0198 2-oxoacid db, 2-oxo acid
 dehydrogenases acyltransferase (catalytic domain), score
 140.90, E-value 2.2e-38; Pfam match to entry PFC0364
 biotin_lipoyl, Biotin-requiring enzyme, score 46.20,
 E-value 7.4e-10
 Similar to Arabidopsis thaliana branched chain alpha-keto

Query Match 21.3%; Score 53; DB 3; Length 158548;
 Best Local Similarity 52.5%; Pred. No. 2.4;
 Matches 116; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 22 AAAAAACACGCTAGAAATCAGATTAAAAATCAGAAATATATCAATTATTCCTG 81
 DB 118893 AAAAAAAAAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTA 118824
 QY 82 ACAATACATGTTCTTTTATACAAAAAATATACACCAATATTAATGTTGTA 141
 DB 118893 AAAAAATATATCCCTCAATATATAAAAAAATTAAGAAATATATATTAATCCATT 118764
 QY 142 TTAGGTATATATATATATATATCTAGGGTTTAACGCAACAAATTATTCGATATA 201
 DB 118763 AT 118704
 QY 202 AATACTTTTAAAGCAACAAATTATATGACGCGAGGAAA 242
 DB 118703 ATGGCTGTACAAAGCAAGTTAAAAAATGGGTAAACAA 118663

RESULT 11
 AC005506/c 205429 bp DNA linear HTG 12-AUG-2000
 LOCUS Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
 DEFINITION
 PROGRESS ***, 4 unordered pieces.
 AC005506
 AC005506.8 GI:9797717

HTG, HTGS, PHASE1.
 SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tanaki, T.,
 Kurdil, O.B., Conway, A.B. and Davis, R.W.
 TITLE Plasmodium falciparum 3D7 chromosome 12
 JOURNAL Unpublished
 2 (bases 1 to 205429)
 Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
 REFERENCE
 AUTHORS Submitted
 TITLE Submitted
 JOURNAL Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA
 94304, USA

COMMENT
 On Aug 12, 2000 this sequence version replaced gi:8810449.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES
 source
 1..205429
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5833"
 /chromosome="12"
 /clone="PFVAC357"
 /clone="3D7"

ORIGIN
 Query Match 21.3%; Score 53; DB 2; Length 205429;
 Best Local Similarity 55.0%; Pred. No. 2.3;
 Matches 104; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 22 AAAAAACACGCTAGAAATCAGATTAAAAATCAGAAATATATCAATTATTCCTG 81
 DB 83338 AATACAAAT 83279
 QY 82 ACAATACATGTTCTTTTATACAAAAAATATACACCAATATTAATGTTGTA 141
 DB 83278 AAT 83219
 QY 142 TTAGGTATATATATATATATATCTAGGGTTTAACGCAACAAATTATTCGATATA 201
 DB 83218 TTAATAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATA 83159
 QY 202 AATACTTTT 210
 DB 83158 TTTT TTTT 83150

RESULT 12
 AE014846 253132 bp DNA linear INV 11-FEB-2003
 LOCUS Plasmodium falciparum 3D7 chromosome 12, section 3 of 9 of the
 DEFINITION complete sequence.
 AC005506 AE014846 AE014188
 VERSION AE014846.1 GI:23496645
 KEYWORDS
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 253132)
AUTHORS Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M.S., Kene V., Shalimov S.J., Sun B., Peterson J., Angiolini S., Perlea M., Allen U., Selengut U., Haft D., Mather M.W., Valdivia A.B., Martin D.M., Falzfa A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M. and Barrall B.
TITLE Genome sequence of the human malaria parasite *Plasmodium falciparum*
JOURNAL Nature 419 (6906), 498-511 (2002)
MEDLINE 2225705
PUBMED 12368864
REFERENCE 2 (bases 1 to 253132)
AUTHORS Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M., Nakao B., Rowley D., Tamaki T., Wang F. and Davis R.W.
TITLE Direct Submission
JOURNAL Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
3 (bases 1 to 253132)
AUTHORS Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M., Nakao B., Rowley D., Tamaki T., Wang F. and Davis R.W.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES
source 1. 253132
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="12"
repeat_region 136..182
/rpt_type=tandem
repeat_region 328..838
/rpt_type=tandem
repeat_region 1400..1425
/rpt_type=tandem
repeat_region 1556..1563
/rpt_type=tandem
repeat_region 1994..2022
/rpt_type=tandem
repeat_region 2249..2282
/rpt_type=tandem
repeat_region 2553..2581
/rpt_type=tandem
repeat_region 2885..2909
/rpt_type=tandem
repeat_region 2970..3091
/rpt_type=tandem
repeat_region 3136..3162
/rpt_type=tandem
gene <3376..>4694
/locus_tag="PFL0565w"
/join(<3376..3481,3717..3818,3990..4253,4432..>4694)
/locus_tag="PFL0565w"
join(3176..3481,3717..3818,3990..4253,4432..4694)
/locus_tag="PFL0565w"
/note="HMPfam hit to PF00226, DnaJ domain, score 2e-39"
/codon_start=1
/product="heat shock protein DnaJ homologue PfJ4"
/protein_id="AN36202.1"
/db_xref="GI:23496646"
/translation="MSRRVYEVGLVPODADLTVTKSYRTLAMKMHDPKRNKKA
ATRPQISRAVEVLSIDPKRRKRYDVGTDENYMDENDEFGNPKPSPNADQILFE
MFGDSSPSPNDSPFSDVWSSFDVDRGRVPSNDPFPNFGSSFNVSFGSSPNDM
DGGSCFTSVSTISNGKFKNRVVKTSKTSIINGKAVRIETVLTILNGTVEVTV
TERREDRGNINROLPAHELNRKR"
repeat_region 3500..3544
/rpt_type=tandem
repeat_region 3873..3905

repeat_region 3917..3956
/rpt_type=tandem
repeat_region 4817..4887
/rpt_type=tandem
repeat_region 4820..4874
/rpt_type=tandem
repeat_region 4936..4994
/rpt_type=tandem
repeat_region 5108..5144
/rpt_type=tandem
repeat_region 5319..5381
/rpt_type=tandem
repeat_region 5565..5636
/rpt_type=tandem
gene complement(<5840..>7003)
/locus_tag="PFL0570c"
complement(<5840..>7003)
/locus_tag="PFL0570c"
complement(5840..7003)
/locus_tag="PFL0570c"
/codon_start=1
/product="hypothetical protein"
/protein_id="AN36203.1"
/db_xref="GI:23496647"
/translation="MTYLAKNKLKNTLRLPFGCNKLVPMRGVINEQKGIIEKT
TDNEKRLNKLKIDVAVKSCNSALPDBCELELYRMKKVDKREVEYDNPSTIEEL
NRELKTRNIPERINRKKGVSNLSYKDIYDPVNPFPDVLNKOELIYENRYSK
MASIVELKQKRAIKKSLKQVLYNPYSEBYENKLNLDNDVKSQOQWCHQDEKRY
MPSFNKRLINKISPIWRYTHLHFIENGILLPRKINMTTRQOQIOPISICIA
RRMALVYDKKPEKEDLIPMDPLQFLDELIRHYIQLNDLRQALIKWIMNYPISLN
VYKFFESKKEKDKHKTQNGENSVSKKPPSQMLQYKQNTYENRYK"
7326..7363
/rpt_type=tandem
repeat_region 7629..7664
/rpt_type=tandem
repeat_region 7668..7702
/rpt_type=tandem
repeat_region 8021..8048
/rpt_type=tandem
repeat_region 8092..8133
/rpt_type=tandem
repeat_region 8274..8311
/rpt_type=tandem
gene <8355..>17469
/locus_tag="PFL0575w"
join(<8355..17080,17320..>17469)
/locus_tag="PFL0575w"
join(8355..17090,17320..17469)
/locus_tag="PFL0575w"
/note="HMPfam hit to PF01593, Flavin containing amine
oxidase, score 1e-05"
/codon_start=1
/product="hypothetical protein"
/protein_id="AN36204.1"
/db_xref="GI:23496648"
/translation="MGTEKEDDQDQIVNRNDLKKEKKNKCDNMSNDISDSISNNIS
DSISTNINNNISNNISNNISNNISNNISNNISNNISNNISNNISNNISNNISNNIS
ISONNHDIIKKVHKNLNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN
YDIPSYLTKSEKSLKQSMIDKNNSEDFPNNVSOKKKNIDILMNNKLTYI
KPSDDIHLIPYRGRFPFTILIGNNILLAKYETVLTLSYLKERTIQCFPSN
FTIDFLVHNNPNNKRYIRNIVINFPLOANNKGDHKKKCTNNNNKKKYITHE
DINNVEYVHDLBNQYSSLPEKCYKKNIKOMNPEKSHNNKIQVEDQKNT
ENKILYLKSKKKINKNNKKNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN
NIVKFNVDVINDVSNKNNHNSLQENNNYVLQIPNNKQIINNIRANVILNNEF
NLQVQPRNTEETNNKIESKNKDEBELCEKKEKQILKEKKLRENOEKRONOEK
ENKKEKILIKKELLILEKILIRKKKEKKLKEKKEKKEKKEFLNNEKEIKKIED
KKRKEPFLAKIYKREKGLQBEKKLQKILINSIKERKKEFLKSKPKKISKKDHL
NKDKTKKQNKIDSQSNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN
DLSISIKKKCKTKNNLLNSINDISDSLSFISTPFPNNIKEDQEDREINDILKL
IEIKVNVQEKNDYIECLHKSXKXVNRKDKIKQOLSKNNKNIITNNPNSNDQ
SDPLNQINSLINDPLNNKNTIIKKYKGGTSVNDMSVMSVYDLSNDVPSNKD
ISNDSIVDISKKNVYVHTNNSLINIDNIHYNNKNNNSERHNSIEVELILKGLKSGF

DIKULSLQKIRIKESKGNKIEYTHVPHNVIESVAKKIVLNKHEIKRKOVCYE
KATKKKISLPGIVKPEDILIKGINNVANIENQULYDVLIIGAGISGLAASYLN
KNKAPLVEGRNIGRAFSSTLPERIINNKLPTVVDLKNVLCNDNDLNDK
KGTKEHPANCTEFNNKSKREDIDYBENFMNKKKKKKKKYVYKRYKNY
KONGDNKGNFISYDVOKEVYTNLFEESKLIINFAPLPEYELNJOSEH
ODENKTKGNKISGNKNIRYLNKNKISCSNINNOYRYSPEDEYSRIRNHR
KILRNNMLPMIKNNDIRNFKIYNKKKNNIPEFSLDIBFVNLSDCVHT
SSLYPRKNTCKSNKLKILSPNKDTRRREDKSVTOLAEKLRVSLVCGKNWE
STVYAHYNNENGRKIKSFKIRVNLCKIRPARRRNPLFGDSNDYQDOK
SEDOQSNNNYSDPDNEXOSDYFNNDILKDKOINNNILDTYHKEILNNHVX
NIKPNENNIIRNENFEASVLYNNKDMKLENEENGKNGKNGEENGKNGKNG
EONBONGKNGKNGEENGKNGKNGKNGKNGKNGKNGKNGKNGKNGKNG
KVDMMNIIISYDHAIIYNNYDGLYNNVRSVTNNIIDEKREKKKEDISKT
EUYNNKRRSMDILMECTEEIPEMNINENFEEMKMLATLOSQYSGSLRE
TSIAMSRLPFSNYLDDMCCKTGSNNIILKIKYIKIKKEOTPYISFKDLNTK
IVDGMKMLINYSIEEONKIFLNTVAELVYIKDKODKKIYHLDNDQIKRQSS
FLNTRKNN
LINNVIKDYKLGENNEDDYKWKIKCCVDYKRNKNGHGTSDILNKNYKVTIYA
KVIYVALPLGCLENNEDDYKWKIKCCVDYKRNKNGHGTSDILNKNYKVTIYA
FTWPSLOINCIIDOKFOPNIIHAGIKGIIIVHCPPKSCYGYKKEHYIINCY
TLKMRKSKKPIPLVDYIITKMODNBSFGSYNPIYNCNDNIIYRSPHPINP
RVVFCBSYLSKSYQCVDGAYDTGISAADIAIGLHNNDKKIDVYIFPQNK
PFTNIPLEPKKFLGYITDGSDEALTDESSDNNMTNDVINDINSNNIDIN
NTNSNTSSNN
KODKNINHIKNETDALCSYKKEIKSNTYESFTNSIINFEKSNDCYSKSTLL
NLYEYNNKSTDOIHYFDRIKSIDSYKRYKNGSYLHLSKDESDNSNDANNTCTI
ETENYNNKNGDHINDYNDYNDYNDYNDYNDYNDYNDYNDYNDYNDYNDY
INKMINKANDAYFYNNIENITLKNQNHAEYKSYFVLLQTFNSIFPKYSPTTFIO
KELPLKIQAQKHLFNDNNFQSLQDNNVVSQLEHNDHVSNNNNNDNNKKKAYE
NTDIINEDEKELSYFHKMEVNNNNKMLSLKSNVANNNDNDNNNNNGHIOQS
NISTDEILNKINLSKTKNNITNKEDVITQTNLIYNNIDIONKIKIYNNKIK
YLNIDYKLYVILHLSISKILDEQISSNNNNKNGSGLFPLCYILOYLKHLIDFN
STVNTKVIYQLCDYIYLFCKHDIKCYKNGGELMNCDSYCNMNSHTYCLASSE
HNNNNKNDKMLCPSCSNIDISKHVGRCYCNONEIMENYWKRIYIKIKFLSLRTR
IRKRLDLKRAHLSRVA"

repeat_region 8427..8510
/rpt_type=tandem
repeat_region 8428..8498
/rpt_type=tandem
repeat_region 8510..8652
/rpt_type=tandem
repeat_region 10256..10314
/rpt_type=tandem
repeat_region 11019..11081
/rpt_type=tandem

Query Match 21.3%; Score 53; DB 3; Length 253132;

Best Local Similarity 55.0%; Pred. No. 2.2; Matches 104; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 22 AAAAAACACGCTAAGAAAATCAGAGTTAAAAAATCAGAAAATATATGATTATTCCTTG 81
DB 71264 ATATACAAAATAATAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 71323
QY 82 ACACATACATGTTCTTTTATACAAAAATATACACACATATTAATGTGTGTA 141
DB 71324 ATATATCATTTTATTTTCAAAAGATATATAAATTTAAAAAATATATATAATAATA 71383
QY 142 TTAGTGTATATTAATTAATTAATCAGAGGTTTAAACGACACATTTATCGATTAATA 201
DB 71384 TTAATAAAAAAAAAAAT 71443
QY 202 AATACTTTT 210
DB 71444 TTTTTTTTT 71452

RESULT 13
PRU251061 609 bp DNA linear PLN 19-MAY-2000
LOCUS PRU251061
DEFINITION Plagioclasma rupestre rps4 gene for small ribosomal protein 4.
ACCESSION AJ251061
VERSION AJ251061.1 GI:8017978

KEYWORDS rps4 gene; small ribosomal protein 4.
SOURCE Plagioclasma rupestre
ORGANISM Plagioclasma rupestre
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Marchantiophyta; Marchantiopsida; Marchantiales;
Marchantiales; Acontaceae; Plagioclasma.
REFERENCE 1
Capeus, I. and Blocher, R.
A molecular approach to bryophyte systematics
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 609)
Capeus, I.
Direct Submission
JOURNAL Submitted (18-NOV-1999) Capeus I., Botanisches Institut,
Universitaet Heidelberg, Im Neuenheimer Feld 360, 69120 Heidelberg,
GERMANY

FEATURES
source
1..609
/organism="Plagioclasma rupestre"
/mol_type="genomic DNA"
/db_xref="taxon:53031"
/tisue_type="gametophyte"
1..609
/gene="rps4"
1..609
/gene="rps4"
/codon_start=1
/product="small ribosomal protein 4"
/protein_id="CAB92154.1"
/db_xref="GI:8017979"
/db_xref="GOA:Q9M4C3"
/db_xref="SWISS-PROT:Q9M4C3"
/translation="MSRYGRPVYIRRLALPGLINPKFRSGYINOSTNKKVQ
YRIERKQRLRFHYGLTERQLKYVIAKAGSTQVQLLEPMDNIIIRLQWA
PTIGARQLVNHQHTLNNNTVDIPSNCKPKDIIITKDKOSQSIITKQINFPQKQ
IPNHLFDLIQFGLINQIVDRECTYKINELLVERYSROV"

ORIGIN

Query Match 21.2%; Score 52.8; DB 8; Length 609;
Best Local Similarity 56.1%; Pred. No. 8.9; Matches 93; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 50 TAAAAATCAGAAAATATATGATTATTCCTTGACACATACATGTTCTTTTATACAA 109
DB 416 TAAAGATCGACAAAATCTCAATCAATATATACAAAATTTAAATTTTCAAAAC 475
QY 110 AATATATACACATATATATGTTGATAGTATAGTATATATATATATAGTAG 169
DB 476 AAAAAATACCAATCATTTTATATATATATATATATATATATATATATATAT 535
QY 170 GGGTTTAAACGACATTTATCGATTAATAATATCTTT 209
DB 536 TTGTTGATCGAATGATTTATTTAAAAATTAATGATT 575

RESULT 14
AF402569 1018 bp DNA linear PLN 05-NOV-2001
LOCUS AF402569
DEFINITION Clavija spinosa ribosomal protein l16 (rpl16) gene, intron
sequence; chloroplast gene for chloroplast product.
ACCESSION AF402569
VERSION AF402569.1 GI:16660274
SOURCE chloroplast Clavija spinosa
ORGANISM Clavija spinosa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; Ericales; Theophrastaceae; Clavija.
REFERENCE 1 (bases 1 to 1018)
Mast, A.R., Kelso, S., Richards, A.J., Lang, D.J., Fellner, D.M.S. and
Conti, E.
Phylogenetic Relationships in Primula L. and Related Genera
(Primulaceae) based on Noncoding Chloroplast DNA

JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL

source

gene

intron

ORIGIN

Query Match 20.9%; Score 52; DB 8; Length 1018;
Best Local Similarity 59.5%; Pred. No. 11;
Matches 88; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Dy 17 TTCGAAAAAACCACGGCTAAAGAAAATCAGAGTTAAAAATCACAAAATATATCATATT 76
TTGGGGGCATAATACATATGATATGATTAAGAAGAAAAAACAAGATTATATTC 716

OY 77 CCTGACACATACATGTCTTTTTTATACAAAAATATATCAACACATATTAATGTG 136
717 TATATCTTANACTATCTATCTTATCTTATTAATTAATTTATAAAAAATATATAT 776

OY 137 TTGTATTAGTGTTATATATAATATATA 164
Dd 777 GTTAATAGTTTAGTTATATATATCTTAA 804

RESULT 15
ACI15683/c

LOCUS

DEFINITION Dictyoestelium discoideum chromosome 2 map complement (6135149-6105997) strain AX4, complete sequence.

ACCESSION ACI15683

VERSION ACI15683.2 GI:28828561

KEYWORDS HTg.

SOURCE Dictyoestelium discoideum

ORGANISM Dictyoestelium discoideum

REFERENCE Gloeckner,G., Eichinger,L., Safranek,K., Pachebett,J., Dear,P., Lehmann,R., Baumgart,C., Farris,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,W., Rosenthal,A. and Noegel,A.A.
Sequence and analysis of chromosome 2 of Dictyoestelium discoideum Nature 418 (6893), 79-85 (2002)

AUTHORS

TITLE The Dictyoestelium Genome Sequencing Consortium

JOURNAL MEDLINE PubMed 12097910

REMARK 2 (bases 1 to 29150)
Baumgart,C.
Direct Submission

REFERENCE Submitted (22-MAR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

JOURNAL 3 (bases 1 to 29150)
Baumgart,C.
Direct Submission

AUTHORS Submitted (04-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE 4 (bases 1 to 29150)
Baumgart,C.
Direct Submission

TITLE Submitted (12-MAR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

JOURNAL

[illegible]

VFKNKFIENIEFNKANNQWIEDEBNKYVNRCKFKYTHSLQIMIKNDLPLIKDK
IKGDIIEIDRSIKYLFSEKSKTPVQYQIETISININNANNKXKEDYL
ELIELMKRDEBEITDILKMAVITNSPDVILNVEPYSVILPTEFEELISNP
ATIKELINLESINFKYGGKLINEPFRKSLAAREKTLKTIETIETIENPISLIP
PLSLINNOKDEBEGIKYVIEPEEIGKSSSLKSLMEPDIVYKRYKTVNFK
MHPYERLKVNIQILIKYCTPSLDTKEIKETIDKIDNIDKDSNHLKLOELALQIF
NQNRFCTYSKTKGRDESKQVNGIQOOLDTYSDEWELNVEEBEENQKOEYF
NNENFRKATVNLSEYSSVYSILYNTMKEGIDITMKLDIKIKYDASGVPELVD
SRFHEQLPEKEVEMKVDYLNESHFKQIGNNYELIMEFESSVTDILSPDLKSNL
EDFMARSLTNSLTNNLNVFOLLNSFDKFNELGVVEMKLNKEKELIEGSD
GKIQOLDKYINQILKSIEMAVEKNTSEIENGVSFSLIINHLLDOLIRIGIS
VLOLDYRSLTPOCLAEKQVYRIFRYNTRSRCLTKYILSRPVSLFTGSRNHS
LVNYTSBSSTLDDKIKLDTFTDYTLGENTSSSGGAGVANNNNNNKFLTD
FTILKELIKQSLDICTVSNIDELDSDEBFWNSFNDKXSLDGRKDFWELLD
WIQVMIKNNEMRYQVQNSPYESLPPSTIFSKVKIELPPTPTTIVAKET
TTIGSRILKIDNFENPTNBSVKIGWENKLEIEFKMSKILNSPFFOREVIRIC
RVANNVEIORFQHNPFNSPHILYSCVANNRDDIMSFLNNYIKENTEPISSE
EKHSQYCEKFLIENHEPEKVNISKLKLEMEREDLOPLVLANAFVPEPCONNIS
TYNKTIONEDLEIRNOAKETIKTLMGEDRENTYNNENFVSVIKENHKTILK
SILEQSDTHLFRKSLDQVLEIKRKTKDEINSLPIECGFSQAFKPNQDIXTWL
ILKTYPNQKINKDSITTEALQONIHILKYKVCNCKNLINNFNDNNLKLKLNEL
SKHPYKXHLTWL
/join(13958, 14003, 14169, 14258, 14352, 14464)
/note="Geneid exon scores (in order of location ranges):
-2.98, 3.35, 6.90 - GSCU_ID dd_00732"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA051170.1"
/db_xref="GI:28828564"
/translation="MGLYRGNCSNRYNSDSIIKSCFTVSGPLDNQMLGSPYEIG
TLIVSTTTTATTTATTAATTAOLQOOLQHNNSYNITW"
complement(15363, 19070)
/note="Geneid exon scores (in order of location ranges):
215.42 - GSCU_ID dd_00730"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA092696.2"
/db_xref="GI:28828565"
/translation="MKNLPLIKDKIKRGGHITTLASIKDLPFSKSTEBPLIKNS
NNKXKEDYELIELMKRREDQETIDLMAVAVSPDITRLVNEPVSIIYPT
TEFALTDSNLEITIKELVNLBPINFRNKGIKILSEPFKRSLSLAHRTKSTSEY
IRNPSLVSIPPLSLINNOQKQKSLNKKLPLIYNNNNNNNDNDEBEIYVID
FEDFKELKSDLSIGLMEIDIVLRTYSGSFIMYHPIYVGLKIKIDIRI
PEKREKILSEKANKIDRNNEGTFNKLKELKLVATEISQINFCYLYWKAARENKIL
CIIHRKGLPDNNIIDEQALVLANHPOAGDEIDEVIELSYNNSDEBEIDIKVVE
QNSPEEYFNYSKRLKESVPPKSNRYSDLNSYLAKEGNIKLEDEYDTNNPQL
DYSKHAQVLPNEKEDVKIWEIPNHSFKQMIQVYGLIEFISVINIVIPLEKIS
NIDDLPSPEKLGTLFQHLNKNFVIELNNSPDINFXQPLEASLIDQERFL
DNEGELQVDEKFAQIIRVIELVMDKVKLTQDKNTENILTSYRLINNLQDL
RYAGLTIDOLEYVRSIIIPQLSNVYEIRIFFYILNRSKLTXYIERPILRYSAY
LNNYDTIAPITNEISCFDLKRFKFDFTYDYLGRRSIQSSGSGSGSGSGS
SSGSGSGSGSKNNKXYITQYDILRELIKOLSHDICMNSRDIYKGDPRLSL
NMDLKSILNIGKDLFWESLWIEQMTKENLMRQFQNSPYESLPTGTIISNV
NIIKELPSKELITLESFLKIDNFBKIPNENSTIKVWNSDKLEPPKIPKLINS
KPSDEVKDSVCVMVNEIORFIQVNFESFPQNEIYVSSCMNREDSMFLNNYNI
EAYFLPGSGYIECKLIFENHEPEKKNISNLMOWESYQSFSALNSFLKYICKE
NNITRYNQIHDNLOIRKQIKELIGLFDRLSCDQNDDEYDDEDDNGEDTN
OSLASIIKQNNKVLLOSIIHOSDTOLFKSALDQFLSEKDEKDEINSPTENN
KNPFOFPKIKNTKODIILEYVNPQKTTKSTIEALKLENIHILKYYKVNCKNL
INNRKDNLLKHLKRELSKHPYKXHLTWL"
/join(20991, 20140, 20295, 20428, 20517, 20611)
/note="Geneid exon scores (in order of location ranges):
-2.09, 13.10, 9.58 - GSCU_ID dd_03401"
/pseudo
/codon_start=1
complement(21212, 21558, 21612, 22326, 22829, 23588,
24212, 24234)
/note="Geneid exon scores (in order of location ranges):
12.40, 25.06, 41.51, -2.31 - GSCU_ID dd_03402"
/codon_start=1
/product="similar to Plasmodium falciparum. Hypothetical
protein"
/protein_id="AA051171.1"

/db_xref="GI:28828566"
/translation="MPIFNKNEPFRKSLAAREKTLKTIETIETIENPISLIP
LSLINNOQKDELSTNRKLSKSTISHNNNNNNNNNNNNNNNNNNNEELKYEIEER
DPKIKNSNLKSFMLDVLSTETNPNQOLMKHVPYRLOVGIKLYKNTIIEK
KREBILIEIKIRKRNDEGNRNLKELRKVLAIEFPQNTFYFYYRQRYRNNIKKLV
FRKRNPFQDNIEFEBHIVQDDTEDEQALVLANHPOAGDEIDEBQVYLFNPSYSLP
NGTIFSNVNIIEKKEFQSKELIETTEAALESFLKIDNEKIPNENSTIKIGNSDKL
IEPPKPSKILNSIFKESDEVISICVMVNEIORFIQVNFESFPQNEIYKYSVMN
RDDIMPLNNYNIIEADFPYFHSQVYCECKLIFENHEPEKKNISNLMOWESY
YLOPPLILNNSPLKYICEKNTITTYNOIONDQOLIRKQVYKTIETIILALDOVLKS
KEPKRDEINSPEIENNNKPIESIFHGDIKQCDIILKYVNPQKTTKDSITTLKMS
NIIHILKYVYDCKNLINNRKDNLLKHLKRELSKHPYKXHLTWL"
complement(24533, 28629)
/note="Geneid exon scores (in order of location ranges):
234.39 - GSCU_ID dd_03403"
/pseudo
/codon_start=1

ORIGIN

Query Match 20.8%; Score 51.8; DB 3; Length 29150;
Best Local Similarity 53.1%; Pred. No. 5.6;
Matches 110; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

CDS

CY 22 AAAAAACGACGCAAGAAATCGATTAAAAATCGAAATATATCATTTCTCTG 81
DB 972 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAACGAAAGATATTTTATTA 913
CY 82 AACATCATGTTCTTTTATTAACAAAAATATATACACATATTAATGTGTGTA 141
DB 912 TGAATACATTTAATCTTTATTAATAAATATATCTTTTATTTTATATCATTTGTT 853
CY 142 TTAGTGTATTAATTAATTAATCTTACGAGGCTTAAACGACCAATTTATCATTAATA 201
DB 852 TAATGTTCTTAATATTTTAAACAGTGCTATTAAGAAATTAATAAAAAAAAAATATAAA 753
CY 202 AATACCTTTAGACGACCAACATTTAT 228
DB 792 AATTAATTAACAAATATATCATTTAT 766

Search completed: February 28, 2004, 07:53:31
Job time: 1257.82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 05:48:29 ; Search time 361.627 Seconds
(without alignments)
2925.120 Million cell updates/sec

Title: US-09-895-435A-4

Perfect score: 1 999gtaccgcgcagcattctg.....gacgcgaggaatcacarg 249

Sequence: 1 999gtaccgcgcagcattctg.....gacgcgaggaatcacarg 249

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	249	100.0	249 7 ABZ75338	ABZ75338 B. thurin
2	244	98.0	244 7 ABZ75335	ABZ75335 B. thurin
3	118	47.4	118 7 ABZ75336	ABZ75336 B. thurin
4	114.4	45.9	116 7 ABZ75340	ABZ75340 B. thurin
5	114.4	45.9	116 7 ABZ75341	ABZ75341 B. thurin
6	64	25.7	116 7 ABZ75342	ABZ75342 B. thurin
7	51	20.5	124 7 ABZ75343	ABZ75343 B. thurin
8	50.2	20.2	6550 6 ABN80270	ABN80270 Human che
9	50	20.1	43545 6 ABK85018	ABK85018 Human cad
10	49.8	20.0	32392 6 ABLS6203	ABLS6203 AnePV gen
11	49.8	20.0	50000 6 ABLS6643	ABLS6643 AnePV gen
12	49	19.7	9905 6 ABLS3062	ABLS3062 Human imm
13	48.8	19.6	9642 6 ABLS3357	ABLS3357 Human imm
14	48.2	19.4	6227 6 ABLS3586	ABLS3586 Human imm
15	48.2	19.4	6227 6 ABLS3586	ABLS3586 Human imm
16	47.4	19.0	17280 4 AAS46772	AAS46772 Chemical
17	47.2	19.0	18585 6 ABLS4609	ABLS4609 Human met
18	46.8	18.8	6072 6 ABLS3030	ABLS3030 Human imm
19	46.8	18.8	6980 6 ABLS3453	ABLS3453 Human imm
20	46.6	18.7	5006 5 AAD2381	AAD2381 Cotton ol
21	46.4	18.6	6012 6 ABK31370	ABK31370 Signal tr
22	46.4	18.6	6012 6 ABLS70327	ABLS70327 Chemical
23	46.4	18.6	6012 6 AAS61274	AAS61274 Human gen

C 24	46.2	18.6	5520 6	ABLS3519	ABLS3519 Human imm
C 25	46	18.5	6328 6	ABN80286	ABN80286 Human che
C 26	45.8	18.4	23683 6	ABLS4623	ABLS4623 Human met
C 27	45.8	18.4	23683 6	ABLS4623	ABLS4623 Human met
C 28	45.6	18.3	6944 7	ABK34027	ABK34027 Human DNA
C 29	45.6	18.3	6944 7	ABK34027	ABK34027 Human DNA
C 30	45.6	18.3	6944 7	ABK34027	ABK34027 Human DNA
C 31	45.4	18.2	3564 6	ABK84093	ABK84093 Human ren
C 32	45.4	18.2	8962 6	ABLS3286	ABLS3286 Human imm
C 33	45.2	18.2	5963 6	ABLS4613	ABLS4613 Human met
C 34	45.2	18.2	6956 6	ABLS70225	ABLS70225 Chemical
C 35	45.2	18.2	9207 6	AAS46790	AAS46790 Tumour su
C 36	45.2	18.2	9207 6	ABK28434	ABK28434 DNA trans
C 37	45.2	18.2	14500 5	ABLS17935	ABLS17935 Human ner
C 38	45.2	18.2	17213 6	ABLS3483	ABLS3483 Human imm
C 39	45.2	18.2	19131 4	AAS46716	AAS46716 Tumour su
C 40	45.2	18.2	23333 4	ABLS4260	ABLS4260 Drosophill
C 41	45	18.1	5875 6	ABLS2288	ABLS2288 Human imm
C 42	45	18.1	6265 2	AAX08523	AAX08523 NBP46 (to
C 43	44.8	18.0	17869 6	ABK39820	ABK39820 Human che
C 44	44.8	18.0	17869 6	ABLS2104	ABLS2104 Human imm
C 45	44.8	18.0	56153 4	AAS46794	AAS46794 Tumour su

ALIGNMENTS

RESULT 1	ABZ75338	standard; DNA; 249 BP.
ID	ABZ75338	
AC	ABZ75338;	
XX		
DT	06-MAY-2003	(first entry)
DE	B. thuringiensis Tnpi recognition target sequence TRT'.	
KW	Tnpi; gene therapy; recombinase; TRT; TRT'; genetic engineering;	
KW	Tnpi recognised target; TRT'; ds.	
XX		
XX	Bacillus thuringiensis.	
OS		
PH	Key	Location/Qualifiers
FT	misc_feature	1..37
FT		/*tag= a
FT		/note= "Tr4430 left IR"
FT	misc_binding	112..145
FT		/*tag= b
FT		/bound_molecy= "Tnpi"
FT	repeat_region	112..143
FT		/*tag= c
FT		/rpt_type= INVERTED
FT	repeat_unit	112..125
FT		/*tag= d
FT	misc_binding	174..193
FT		/*tag= e
FT		/bound_molecy= "Tnpi"
FT	repeat_region	176..227
FT		/*tag= f
FT		/rpt_type= INVERTED
FT	repeat_unit	176..191
FT		/*tag= g
FT	misc_binding	209..229
FT		/*tag= h
FT		/bound_molecy= "Tnpi"
PN	MO2003004652-A2.	
XX		
PD	16-JAN-2003.	
XX		
PF	28-JUN-2002; 2002WO-EP007176.	
XX		
PR	30-JUN-2001; 2001US-00895435.	


```

XX (EURO-) EURO MOLECULAR BIOLOGY LAB.
PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI Stewart FA, Zhang Y, Hallet B;
XX WPI; 2003-221599/21.
XX
PT A composition comprising recombinase TnpI and the recombination substrate
PT TRT or TRT' sequences, useful for genetic engineering, TnpI-mediated
PT genetic recombination, or for gene therapy.
XX
PS Claim 1; Page 55; 58pp; English.
XX
CC The invention relates to a novel composition comprising an isolated DNA
CC molecule comprising TnpI recombination substrates. The composition may
CC have a use in gene therapy, and as a vaccine. The recombinase TnpI and
CC the recombination substrates TRT or TRT' sequences are useful for genetic
CC engineering. The kits are useful for TnpI-mediated genetic recombination.
CC The TnpI-mediated site-specific recombination is useful for deleting or
CC inserting DNA sequences to modulate gene expression, integrating new
CC nucleotide sequences into a DNA molecule, such as an episome or genomic
CC DNA, for gene therapy, or for producing food-grade genetically modified
CC organisms. The present sequence represents the TnpI recognised target
CC (TRT) TRT'.
XX
SQ Sequence 249 BP; 103 A; 38 C; 36 G; 72 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 249; DB 7; Length 249;
Best Local Similarity 100.0%; Pred. No. 7.7e-41;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGGGTACCGCCAGCATTTGGGAAAAAACCAGCTAAGAAATCAGATTAAAAATCAG 60
DB 1 GGGGTACCGCCAGCATTTGGGAAAAAACCAGCTAAGAAATCAGATTAAAAATCAG 60
QY 61 AAAATATATCATTTATCTCTTGACACATACATGTTCTTTTATACAAAAATATACAA 120
DB 61 AAAATATATCATTTATCTCTTGACACATACATGTTCTTTTATACAAAAATATACAA 120
QY 121 CACAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 180
DB 121 CACAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 180
QY 121 CACAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 180
DB 121 CACAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 180
QY 161 AACCAATTTATGATTAATTAATTAATCTTTAGACGACACCAATTTATAGCGCGAGGA 240
DB 161 AACCAATTTATGATTAATTAATTAATCTTTAGACGACACCAATTTATAGCGCGAGGA 240
QY 241 AATCACAATG 249
DB 241 AATCACAATG 249
XX
RESULT 2
ID ABZ75335 standard; DNA; 244 BP.
ABZ75335
AC ABZ75335;
XX
DT 06-MAY-2003 (first entry)
XX
DE B. churingiensis TnpI recognition target sequence TRT'.
XX
KW TnpI; gene therapy; recombinase; TRT; TRT'; genetic engineering;
XX TnpI recognised target; TRT'; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT repeat_region 37..59
FT /*tag= a
FT /rpt_type= DIRECT
FT repeat_unit 37..45

```

```

FT /*tag= b
FT repeat_region 111..142
FT /*tag= c
FT /rpt_type= INVERTED
FT repeat_unit 111..124
FT /*tag= d
FT repeat_region 175..211
FT /*tag= e
FT /rpt_type= DIRECT
FT repeat_unit 175..190
FT /*tag= f
XX
EN WO2003004652-A2.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-EP007176.
XX
PR 30-JUN-2001; 2001US-00895435.
XX
PA (EURO-) EURO MOLECULAR BIOLOGY LAB.
PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
PI Stewart FA, Zhang Y, Hallet B;
XX WPI; 2003-221599/21.
XX
DR WPI; 2003-221599/21.
XX
XX
XX A composition comprising recombinase TnpI and the recombination substrate
XX TRT or TRT' sequences, useful for genetic engineering, TnpI-mediated
XX genetic recombination, or for gene therapy.
XX
PS Claim 6; Fig 1A; 58pp; English.
XX
CC The invention relates to a novel composition comprising an isolated DNA
CC molecule comprising TnpI recombination substrates. The composition may
CC have a use in gene therapy, and as a vaccine. The recombinase TnpI and
CC the recombination substrates TRT or TRT' sequences are useful for genetic
CC engineering. The kits are useful for TnpI-mediated genetic recombination.
CC The TnpI-mediated site-specific recombination is useful for deleting or
CC inserting DNA sequences to modulate gene expression, integrating new
CC nucleotide sequences into a DNA molecule, such as an episome or genomic
CC DNA, for gene therapy, or for producing food-grade genetically modified
CC organisms. The present sequence represents the TnpI recognised target
CC (TRT) TRT'.
XX
SQ Sequence 244 BP; 102 A; 38 C; 33 G; 71 T; 0 U; 0 Other;
XX
Query Match 98.0%; Score 244; DB 7; Length 244;
Best Local Similarity 100.0%; Pred. No. 7.7e-40;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3 GGTACCGCCAGCATTTGGGAAAAAACCAGCTAAGAAATCAGATTAAAAATCAGAA 62
DB 1 GGTACCGCCAGCATTTGGGAAAAAACCAGCTAAGAAATCAGATTAAAAATCAGAA 60
QY 63 AATATATCATTTATCTCTTGACACATACATGTTCTTTTATACAAAAATATACAA 122
DB 63 AATATATCATTTATCTCTTGACACATACATGTTCTTTTATACAAAAATATACAA 120
QY 123 CAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 182
DB 123 CAATATTAATGTGTGTATTAAGTGTATTAATAATTAATTAATCTAGGGGTTTAAACG 180
QY 183 CACAATTTATGATTAATTAATTAATCTTTAGACGACACCAATTTATAGCGCGAGGAAA 242
DB 183 CACAATTTATGATTAATTAATTAATCTTTAGACGACACCAATTTATAGCGCGAGGAAA 240
QY 243 TCAC 246
DB 241 TCAC 244
XX
RESULT 3

```

```
AB275336
ID AB275336 standard; DNA; 118 BP.
XX
AC AB275336;
XX
DT 06-MAY-2003 (first entry)
XX
DE B. thuringiensis TnpI recognition target sequence TRT'.
XX
KM TnpI; gene therapy; recombinase; TRT; TRT'; genetic engineering;
KW TnpI recognised target; TRT'; ds.
XX
OS Bacillus thuringiensis.
XX
Key Location/Qualifiers
FH repeat_region 1..32 /*tag= a
FT /*rpt_type= INVERTED
FT repeat_unit 1..14 /*tag= b
FT repeat_region 65..116 /*tag= c
FT /*rpt_type= DIRECT
FT repeat_unit 65..80 /*tag= d
FT /*tag= d

WO2003004652-A2.
XX
XX 16-JAN-2003.
XX
XX 28-JUN-2002; 2002WO-EP007176.
XX
XX 30-JUN-2001; 2001US-00895435.
XX
XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
XX Stewart FA, Zhang Y, Hallet B;
XX
XX WPI; 2003-221599/21.
XX
XX A composition comprising recombinase TnpI and the recombination substrate
XX TRT or TRT' sequences, useful for genetic engineering, TnpI-mediated
XX genetic recombination, or for gene therapy.
XX
XX Claim 6; Page 55; 58pp; English.
XX
XX The invention relates to a novel composition comprising an isolated DNA
XX molecule comprising TnpI recombination substrates. The composition may
XX have a use in gene therapy, and as a vaccine. The recombinase TnpI and
XX the recombination substrates TRT or TRT' sequences are useful for genetic
XX engineering. The kits are useful for TnpI-mediated genetic recombination.
XX The TnpI-mediated site-specific recombination is useful for deleting or
XX inserting DNA sequences to modulate gene expression, integrating new
XX nucleotide sequences into a DNA molecule, such as an episome or genomic
XX DNA, for gene therapy, or for producing food-grade genetically modified
XX organisms. The present sequence represents the TnpI recognised target
XX (TRT) TRT'.
XX
XX Sequence 118 BP; 49 A; 14 C; 15 G; 40 T; 0 U; 0 Other;
SQ
XX
XX Query Match 47.4%; Score 118; DB 7; Length 118;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-15;
XX Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 113 TATATCAACACAAATTAATGTTGATAGGTTATTAATAATAATCTAGGG 172
XX TATATCAACACAAATTAATGTTGATAGGTTATTAATAATAATCTAGGG 60
XX
XX 173 TTAAACGACACAAATTTATGATTAATAATAATCTAGGACGACACAAATTATAG 230
XX TTAAACGACACAAATTTATGATTAATAATAATCTAGGACGACACAAATTATAG 118
XX
```

```
RESULT 4
AB275341
ID AB275341 standard; DNA; 116 BP.
XX
AC AB275341;
XX
DT 06-MAY-2003 (first entry)
XX
DE B. thuringiensis TnpI recognition target TRT' functional variant #2.
XX
KM TnpI; gene therapy; recombinase; TRT; TRT'; genetic engineering;
KW TnpI recognised target; TRT'; ds.
XX
OS Bacillus thuringiensis.
XX
Key Location/Qualifiers
FH repeat_region 1..32 /*tag= a
FT /*rpt_type= INVERTED
FT repeat_unit 1..14 /*tag= b
FT repeat_region 65..116 /*tag= c
FT /*rpt_type= DIRECT
FT repeat_unit 65..101 /*tag= d
FT /*tag= d

WO2003004652-A2.
XX
XX 16-JAN-2003.
XX
XX 28-JUN-2002; 2002WO-EP007176.
XX
XX 30-JUN-2001; 2001US-00895435.
XX
XX (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
XX Stewart FA, Zhang Y, Hallet B;
XX
XX WPI; 2003-221599/21.
XX
XX A composition comprising recombinase TnpI and the recombination substrate
XX TRT or TRT' sequences, useful for genetic engineering, TnpI-mediated
XX genetic recombination, or for gene therapy.
XX
XX Disclosure; Page 12; 58pp; English.
XX
XX The invention relates to a novel composition comprising an isolated DNA
XX molecule comprising TnpI recombination substrates. The composition may
XX have a use in gene therapy, and as a vaccine. The recombinase TnpI and
XX the recombination substrates TRT or TRT' sequences are useful for genetic
XX engineering. The kits are useful for TnpI-mediated genetic recombination.
XX The TnpI-mediated site-specific recombination is useful for deleting or
XX inserting DNA sequences to modulate gene expression, integrating new
XX nucleotide sequences into a DNA molecule, such as an episome or genomic
XX DNA, for gene therapy, or for producing food-grade genetically modified
XX organisms. The present sequence represents a TnpI recognised target (TRT)
XX functional variant TRT'.
XX
XX Sequence 116 BP; 47 A; 14 C; 14 G; 41 T; 0 U; 0 Other;
SQ
XX
XX Query Match 45.9%; Score 114.4; DB 7; Length 116;
XX Best Local Similarity 99.1%; Pred. No. 5e-14;
XX Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 113 TATATCAACACAAATTAATGTTGATAGGTTATTAATAATAATCTAGGG 172
XX TATATCAACACAAATTAATGTTGATAGGTTATTAATAATAATCTAGGG 60
XX
XX 173 TTAAACGACACAAATTTATGATTAATAATAATCTAGGACGACACAAATTAT 228
XX TTAAACGACACAAATTTATGATTAATAATAATCTAGGACGACACAAATTAT 118
XX
```

DB 61 TTTAAGCAACACATTATTCGATAATAATTAATCTTTAGACGACACAAATTAT 116
RESULT 5
ID ABZ75340 standard; DNA; 116 BP.
AC ABZ75340;
XX
DT 06-MAY-2003 (first entry)
XX
DE B. thuringiensis Tnpi recognition target TRT' functional variant #1.
XX
DE Tnpi; gene therapy; recombinase; TRT'; TRT'; genetic engineering;
XX
KM Tnpi recognised target; TRT'; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FH repeat_region 1..32
FT /*tag= a
FT repeat_unit 1..14
FT /*tag= b
FT repeat_region 65..116
FT /*tag= c
FT /*tag= c
FT /*tag= c
FT repeat_unit 65..80
FT /*tag= d
XX
PN MO2003004652-A2.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002MO-EP007176.
XX
PR 30-JUN-2001; 2001US-00895435.
XX
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI Stewart FA, Zhang Y, Hallet B;
XX
DR WPI; 2003-221599/21.
XX
PT A composition comprising recombinase Tnpi and the recombination substrate
PT TRT' or TRT' sequences, useful for genetic engineering, Tnpi-mediated
PT genetic recombination, or for gene therapy.
XX
PS Disclosure; Page 12; 58pp; English.
XX
CC The invention relates to a novel composition comprising an isolated DNA
CC molecule comprising Tnpi recombination substrates. The composition may
CC have a use in gene therapy, and as a vaccine. The recombinase Tnpi and
CC the recombination substrates TRT' or TRT' sequences are useful for genetic
CC engineering. The kits are useful for Tnpi-mediated genetic recombination.
CC The Tnpi-mediated site-specific recombination is useful for deleting or
CC inserting DNA sequences to modulate gene expression, integrating new
CC nucleotide sequences into a DNA molecule, such as an episome or genomic
CC DNA, for gene therapy, or for producing food-grade genetically modified
CC organisms. The present sequence represents a Tnpi recognised target (TRT')
CC functional variant TRT'.
XX
SQ Sequence 116 BP; 47 A; 14 C; 14 G; 41 T; 0 U; 0 Other;
Query Match 45.9%; Score 114.4; DB 7; Length 116;
Best Local Similarity 99.1%; Pred. No. 5e-14;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 113 TAATACACACATATTAATGTTGATTTAGTGTATTAATAATATTAATCTAGGGG 172
DB 1 TAATACACACATATTAATGTTGATTTAGTGTATTAATAATATTAATCTAGGGG 60

OY 173 TTTAAGCAACACATTATTCGATAATAATTAATCTTTAGACGACACAAATTAT 228
DB 61 TTTAAGCAACACATTATTCGATAATAATTAATCTTTAGACGACACAAATTAT 116
RESULT 6
ID ABZ75342 standard; DNA; 116 BP.
AC ABZ75342;
XX
DT 06-MAY-2003 (first entry)
XX
DE B. thuringiensis Tnpi recognition target TRT' functional variant #3.
XX
DE Tnpi; gene therapy; recombinase; TRT'; TRT'; genetic engineering;
XX
KM Tnpi recognised target; TRT'; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FH repeat_region 1..32
FT /*tag= a
FT repeat_unit 1..14
FT /*tag= b
FT repeat_region 65..116
FT /*tag= c
FT /*tag= c
FT repeat_unit 65..80
FT /*tag= d
XX
PN MO2003004652-A2.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002MO-EP007176.
XX
PR 30-JUN-2001; 2001US-00895435.
XX
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI Stewart FA, Zhang Y, Hallet B;
XX
DR WPI; 2003-221599/21.
XX
PT A composition comprising recombinase Tnpi and the recombination substrate
PT TRT' or TRT' sequences, useful for genetic engineering, Tnpi-mediated
PT genetic recombination, or for gene therapy.
XX
PS Disclosure; Page 13; 58pp; English.
XX
CC The invention relates to a novel composition comprising an isolated DNA
CC molecule comprising Tnpi recombination substrates. The composition may
CC have a use in gene therapy, and as a vaccine. The recombinase Tnpi and
CC the recombination substrates TRT' or TRT' sequences are useful for genetic
CC engineering. The kits are useful for Tnpi-mediated genetic recombination.
CC The Tnpi-mediated site-specific recombination is useful for deleting or
CC inserting DNA sequences to modulate gene expression, integrating new
CC nucleotide sequences into a DNA molecule, such as an episome or genomic
CC DNA, for gene therapy, or for producing food-grade genetically modified
CC organisms. The present sequence represents a Tnpi recognised target (TRT')
CC functional variant TRT'.
XX
SQ Sequence 116 BP; 27 A; 11 C; 5 G; 21 T; 0 U; 52 Other;
Query Match 25.7%; Score 64; DB 7; Length 116;
Best Local Similarity 55.2%; Pred. No. 0.00054;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 113 TAATACACACATATTAATGTTGATTTAGTGTATTAATAATATTAATCTAGGGG 172

Dn TAAATACAGCAACAATTATACTGTTGTGCATANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Gy 173 TTTAAGCGAACACACTTTCATATAAATTAACCTTAGCCGAACAACTTAT 228
|||||
Db NNNNACGACACAACTTATNTNNNNNNNNNNNNNNNNNNNNNAACGAACAACATTTAT 116
|||||

RESULT 7
ID ABZ75343
ABZ75343 standard; DNA; 124 BP.

XX AC ABZ75343;
XX DT 06-MAY-2003 (first entry)
XX DE B. thuringiensis Thpi recognition target TRT consensus.
XX KM Thpi; gene therapy; recombinase; TRT; TRT'; genetic engineering;
XX RW Thpi recognised target; TRT'; ds; consensus.
XX OS Bacillus thuringiensis.

Key Location/Qualifiers
FT repeat_region 1..32 /tag= a
FT FT /ipc_type= INVERTED
FT repeat_unit 1..14
FT FT /*tag= b
FT repeat_region 75..124
FT FT /*tag= c
FT repeat_unit /ipc_type= DIRECT
FT FT 75..89
FN repeat_tag d
PN WO2003004652-A2.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002WO-EPO07176.
XX PF
PR 30-JUN-2001; 2001US-00895435.
XX PR
PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
PA (UTLO-) UNIV CATHOLIQUE LOUVAIN.
PI Stewart FA, Zhang Y, Hallett B;
PI PI
XX STW; 2003-221599/21.
DR
XX

A composition comprising recombinase Thpi and the recombination substrate TRT or TRT' sequences, useful for genetic engineering, Thpi-mediated genetic recombination, or for gene therapy.

Disclosure; Page 13; 58pp; English.

The invention relates to a novel composition comprising an isolated DNA molecule comprising Thpi recombination substrates. The composition may have a use in gene therapy, and as a vaccine. The recombinase Thpi and CC the recombination substrates TRT or TRT' sequences are useful for genetic engineering. The kits are useful for Thpi-mediated genetic recombination. CC The Thpi-mediated site-specific recombination is useful for deleting or CC inserting DNA sequences to modulate gene expression, integrating new CC nucleotide sequences into a DNA molecule, such as an episome or genomic CC DNA, for gene therapy, or for producing food-grade genetically modified CC organisms. The present sequence represents a Thpi recognisable target (TRT) CC consenus sequence

SQ Sequence 124 BP; 36 A; 13 G; 26 T; 0 U; 42 Other;

Query Match 20.5%; Score 51; DB 7; Length 124;
Best Local Similarity 65.6%; Freq.No. 0.21;
Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps 2

[illegible]

SQ Sequence 6550 BP; 1589 A; 235 C; 1752 G; 2974 T; 0 U; 0 Other;
Query Match 20.2%; Score 50.2; DB 6; Length 6550;
Best Local Similarity 53.9%; Pred. No. 0.34; Mismatches 89; Indels 0; Gaps 0;
Matches 103; Conservative 0;
QY 23 AAAAACCACGCTAGAGAAATCAGAGTTTAAATAATCAGAAATATATATTCCTTGA 82
DB 687 AACTAAAAACCAACCAATACACCTTTACTTAAACTCAAACCTATATATTCCTTA 628
QY 83 CACATACAGTCTCTTTTATACAAAAATATACACCAATATTAATTTGTTCTAT 142
DB 627 ATATATATATCTCTCTATAAAAAATATATATCTCTTTAAAAACCAATTTAT 569
QY 143 TAGGTGTTATATATATATATATATCTAGGGTTTACGCAACACATTTATCGATAATAA 202
DB 567 ACTTCTCTTTAAAAAATAATTAATTAATAAATAAATAAATTTTCATTAAACAA 508
QY 203 ATACTTTTGA 213
DB 507 ATAATTTAATA 497

RESULT 9
ABK85018/c
ID ABK85018 standard; DNA; 43545 BP.
XX
AC ABK85018;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human cadherin-like asymmetry protein-2 (CLASP-2) genomic DNA.
XX
KW Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW autoimmune pulmonary inflammation; organ rejection; inflammation; CLASP;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200231117-A2.
XX
PD 18-APR-2002.
XX
PF 15-OCT-2001; 2001WO-US032202.
XX
PR 13-OCT-2000; 2000US-00687837.
XX
PA (ARBO-) ARBOR VITA CORP.
PA (GARM/) GARMAN J D.
PA (CAND/) CANDIA A F.
XX
PI Lu PS;
XX
DR WPI; 2002-416861/44.
XX
PT New human cadherin-like asymmetry protein(s) (CLASP)-2 for modulating an
PT immune response, and for treating multiple sclerosis, rheumatoid
PT arthritis, endometriosis, lupus, autoimmune thyroiditis, septic shock,
PT and sepsis.
XX
PS Disclosure; Fig 12B; 245pp; English.
XX
CC The invention relates to an isolated polypeptide (I) comprising an amino
CC acid sequence that has 90 % sequence identity to one of the human
CC cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC sequences (PS). (I) is useful for identifying a compound or agent that
CC binds CLASP-2 polypeptide. An antibody (II) to (I) is useful for
CC detecting a CLASP-2 polypeptide in a sample. (II) is useful for

CC inhibiting a immune response in a subject. A pharmaceutical composition
CC comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC the autoimmune disease is caused or exacerbated by increased activity of
CC TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC detection or inhibition of CLASP-2 expression (e.g., antisense or
CC ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC polynucleotides can express CLASP-2 polypeptides, produce anti-CLASP-
CC polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC purposes. CLASP-2 polynucleotides can construct transgenic and knockout
CC animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
CC polynucleotides or polynucleotides can treat deficiencies or disorders of
CC the immune system, by activating or inhibiting the activation,
CC differentiation of immune cells and can treat or detect deficiencies or
CC disorders of haematopoietic cells. CLASP-2 polypeptides or
CC polynucleotides can increase differentiation and proliferation of
CC haematopoietic cells, including the pluripotent stem cells to treat those
CC disorders associated with a decrease in certain (or many) types of
CC haematopoietic cells e.g., immunologic deficiency syndromes including
CC blood protein disorders (e.g., agammaglobulinaemia,
CC dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC haemoglobinuria). CLASP-2 polynucleotides or polypeptides can treat or
CC detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC inflammation. CLASP-2 can be used to treat anaphylaxis or
CC hypersensitivity to an antigenic molecules, organ rejection or graft-
CC versus-host disease (GVHD) and inflammation. ABK84922-ABK85018 represent
CC cadherin-like asymmetry protein (CLASP) coding sequences and PCR primers
CC of the invention
XX
SQ Sequence 43545 BP; 12383 A; 8378 C; 8983 G; 13801 T; 0 U; 0 Other;

Query Match 20.1%; Score 50; DB 6; Length 43545;
Best Local Similarity 61.5%; Pred. No. 0.39; Mismatches 0; Gaps 0;
Matches 80; Conservative 0;
QY 36 AAGAAAATCAGAGTTTAAAAATCAGAAAATATATATATTCCTTGACACATACATGTC 95
DB 7809 AATAAAAAATAAAAAATAAAAAAGTAATAATTTATCTTCTCCATAAGATAATTTTATT 7750
QY 96 TTTTATATACAAAAATATATACACACATATATTCGTTGTTATAGGTGTATAAT 155
DB 7749 GCATTTCAACAAATAATATATATTAATCAATGATCTTGATAGGATATATAGTAT 7690
QY 156 AAATATAAAT 165
DB 7689 CAATATATAT 7680

RESULT 10
ABL56203/c
ID ABL56203 standard; DNA; 32392 BP.
XX
AC ABL56203;
XX

DT 01-JUL-2002 (first entry)
XX
DE AmEPV genome fragment#5.
XX
KW AmEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
KW genetic deficiency disorder; ds.
XX
OS Amsacta moorei entomopoxvirus.
XX
PN WO200212526-A2.
XX
PD 14-FEB-2002.
XX

ABL32062/c
ID ABL32062 standard; DNA; 9905 BP.
XX AC ABL32062;
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 35.
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;
XX KW antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX OS Homo sapiens.
XX PN WO200200928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP007537.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful for
XX PT diagnosis and treatment of diseases associated with abnormal cytosine
XX PT methylation.
XX PS Claim 1; SEQ ID NO 35; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention
XX SQ Sequence 9905 BP; 3133 A; 75 C; 1920 G; 4777 T; 0 U; 0 Other;
Query Match 19.7%; Score 49; DB 6; Length 9905;
Best Local Similarity 51.1%; Pred. No. 0.6;
Matches 115; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 7 CCGCCAGCATTCGGAAGAAACACCGCTAAGAAATCAGAGTTAAAGAAATCAGAAATA 66
Db 9905 CCTAAGACATCAACAAAGAAATATACCAATATTTCCACGCTTACATCAAAATTA 9846
QY 67 TATCATTTCTTGACATACATGTTCTTTTATACAAAAATATACACACAAAT 126
Db 9845 TATCTTTTCATCTTTATACATACAAATTAATCCCATATCTTAACACCTTAACAAATATTA 9786
QY 127 ATTAATTTGTTCTTATAGTGTATATAATATAATCTAGGGTTTAAACGCAACACA 186
Db 9785 CATAAAAATAAAATCCTTAATTTATAAAAATAATATAATATCTTTTAAAAAACAATA 9726
QY 187 ATTATCGATAAATAAATACTTTTAGACGCAACACAAATTTATAGA 231
Db 9725 ACATAAAAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9681

RESULT 13
ABL32357/c
ID ABL32357 standard; DNA; 9642 BP.
XX AC ABL32357;
XX DT 26-MAR-2002 (first entry)
XX DE Human immune system associated gene SEQ ID NO: 330.
XX DE Human; immune system disease; cytosine methylation; antiasthmatic;
XX KW antiarteriosclerotic; anti-HIV; anticonvulsant; ophthalmological;
XX KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
XX KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX OS Homo sapiens.
XX PN WO200200928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP007537.
XX PR 30-JUN-2000; 2000DE-01032529.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful for
XX PT diagnosis and treatment of diseases associated with abnormal cytosine
XX PT methylation.
XX PS Claim 1; SEQ ID NO 330; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated
XX CC genes which are modified by the methylation of cytosines. The sequences
XX CC can be used in the diagnosis and treatment of immune system disorders,
XX CC including eye diseases such as retinopathy, neovascular glaucoma and
XX CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX CC diseases. The present sequence is a gene of the invention
XX SQ Sequence 9642 BP; 2381 A; 88 C; 2349 G; 4824 T; 0 U; 0 Other;
Query Match 19.6%; Score 48.8; DB 6; Length 9642;
Best Local Similarity 51.4%; Pred. No. 0.65;
Matches 113; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 12 AGCATTTTCGAAAAAACCACGCTAAGAAATCAGAGTTAAAGAAATCAGAAATAATATCA 71
Db 3816 AACATTCGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3757
QY 72 TTATTCCTTGACATACATGTTCTTTTATACAAAAATATACACACAAATTTAA 131
Db 3756 TACCTTAAAAACCAAAAAATAAAAAAATCTTAAAAAATAAAAAAATTTACCAAAATATAA 3697
QY 132 TTGTGTTGTTAGTGTATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 191
Db 3696 ATATAAAAAAACCCCTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3637
QY 192 TCGATAAATAAATACTTTTAGACGCAACACAAATTTATAGA 231
Db 3636 ACAAAATATACATCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3597

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:07:41 ; Search time 274.782 Seconds
(without alignments)
3268.988 Million cell updates/sec

Title: US-09-895-435A-4

Perfect score: 249
Sequence: 1 ggggtaccgcagcatttcg.....gacggagggaatcacatg 249

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2353733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707465

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US03_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	100.0	249	10	US-09-895-435-4
2	244	98.0	244	10	US-09-895-435-1
3	118	47.4	118	10	US-09-895-435-2
4	114.4	45.9	116	10	US-09-895-435-7
5	114.4	45.9	116	10	US-09-895-435-8
6	64	25.7	116	10	US-09-895-435-9
7	51	20.5	124	10	US-09-895-435-10
8	50.2	20.2	761	12	US-10-424-599-137579
9	49	19.7	9905	14	US-10-311-455-35
10	48.8	19.6	9642	14	US-10-311-455-330
11	48.2	19.4	6227	14	US-10-311-455-1559
12	47.2	19.0	18585	14	US-10-240-485-162
13	47.2	19.0	367378	14	US-10-312-841-1
14	46.8	18.8	6072	14	US-10-311-455-3
15	46.8	18.8	6980	14	US-10-311-455-426

C 16	46.6	18.7	5006	9	US-09-837-751-7	Sequence 7, Appli
C 17	46.6	18.7	367378	14	US-10-312-841-2	Sequence 2, Appli
C 18	46.4	18.6	6012	12	US-10-221-613-235	Sequence 235, App
C 19	46.4	18.6	367378	14	US-10-312-841-1	Sequence 1, Appli
C 20	46.2	18.6	5520	14	US-10-311-455-1492	Sequence 1492, Ap
C 21	45.8	18.4	23683	14	US-10-240-485-176	Sequence 176, App
C 22	45.6	18.3	6944	14	US-10-172-086-112	Sequence 112, App
C 23	45.4	18.2	8962	14	US-10-311-455-659	Sequence 659, App
C 24	45.2	18.2	434	12	US-10-424-599-86977	Sequence 86977, A
C 25	45.2	18.2	593	14	US-10-240-485-166	Sequence 166, App
C 26	45.2	18.2	9207	14	US-10-240-453-308	Sequence 308, App
C 27	45.2	18.2	17213	14	US-10-311-455-1456	Sequence 1456, Ap
C 28	45	18.1	5875	14	US-10-311-455-261	Sequence 261, App
C 29	45	18.0	6265	9	US-09-129-112-3	Sequence 3, Appli
C 30	44.8	18.0	17869	14	US-10-311-455-77	Sequence 77, Appl
C 31	44.8	18.0	17869	16	US-10-257-166-1	Sequence 1, Appli
C 32	44.8	18.0	113515	14	US-10-311-455-2148	Sequence 2148, Ap
C 33	44.6	17.9	11790	14	US-10-311-455-516	Sequence 516, App
C 34	44.6	17.9	17594	14	US-10-311-455-2000	Sequence 2000, Ap
C 35	44.6	17.9	19659	14	US-10-311-455-740	Sequence 740, App
C 36	44.4	17.8	5621	14	US-10-311-455-177	Sequence 177, App
C 37	44.4	17.8	6782	14	US-10-311-455-749	Sequence 749, App
C 38	44.4	17.8	17594	14	US-10-311-455-1999	Sequence 1999, Ap
C 39	44.2	17.8	431	9	US-09-960-352-5558	Sequence 5558, Ap
C 40	44.2	17.8	5532	14	US-10-311-455-752	Sequence 752, App
C 41	44.2	17.8	6154	12	US-10-221-613-69	Sequence 69, Appl
C 42	44.2	17.8	13326	14	US-10-311-455-1832	Sequence 1832, Ap
C 43	44	17.7	5689	14	US-10-239-676-90	Sequence 90, Appl
C 44	44	17.7	5689	14	US-10-240-453-100	Sequence 100, App
C 45	44	17.7	6223	12	US-10-221-613-136	Sequence 136, App

ALIGNMENTS

RESULT 1

US-09-895-435-4
; Sequence 4, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-4

Query Match	100.0%;	Score 249;	DB 10;	Length 249;
Best Local Similarity	100.0%;	Pred. No. 2.6e-40;		
Matches 249;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGGTACCGCAGCATTTCCGAAAAAACCACGCTAAGAAATCAGATTAATAATTCAG	60	
Db	1	GGGTACCGCAGCATTTCCGAAAAAACCACGCTAAGAAATCAGATTAATAATTCAG	60	
Qy	61	AAATATATCATTTATTCCTTGACACATACATGTTCTTTTATACAAAAATAATACAA	120	
Db	61	AAATATATCATTTATTCCTTGACACATACATGTTCTTTTATACAAAAATAATACAA	120	
Qy	121	CACAATTAATTTGTTGTTAGGTTTATTAATAATAATCTAGGGTTTAAACGC	180	
Db	121	CACAATTAATTTGTTGTTAGGTTTATTAATAATAATCTAGGGTTTAAACGC	180	
Qy	181	AACCAATTTATCGATAAATAAATCTTTTAGCGGACACAAATTTATAGCCGGAGGA	240	

Db 181 AACACATTTATCGATAAATAAATCTTTTAGACGCAACACAAATTTATAGACGGGAGGA 240
QY 241 AATCACATG 249
Db 241 AATCACATG 249

RESULT 2
US-09-895-435-1
; Sequence 1, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 244
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-1

Query Match 98.0%; Score 244; DB 10; Length 244;
Best Local Similarity 100.0%; Pred. No. 2.6e-39;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GGTACGCCGAGCATTCGGAAGAAACACCGCTAGAAATCAGAGTTAAATAATCAGAA 62
Db 1 GGTACGCCGAGCATTCGGAAGAAACACCGCTAGAAATCAGAGTTAAATAATCAGAA 60
QY 63 AATATATCATTTCTTCGACACATACATGTTCTTTTATACAAAAATAATACACA 122
Db 61 AATATATCATTTCTTCGACACATACATGTTCTTTTATACAAAAATAATACACA 120
QY 123 CAATATTAATGTTGTTAGTTAGTTATTAATAATAATCTAGGGTTTACGCA 182
Db 121 CAATATTAATGTTGTTAGTTAGTTATTAATAATAATCTAGGGTTTACGCA 180
QY 183 CACAATTTATCGATAAATAAATCTTTTAGACGCAACACAAATTTATAGCGGGAGGAA 242
Db 181 CACAATTTATCGATAAATAAATCTTTTAGACGCAACACAAATTTATAGCGGGAGGAA 240
QY 243 TCAC 246
Db 241 TCAC 244

RESULT 3
US-09-895-435-2
; Sequence 2, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-2
Query Match 47.4%; Score 118; DB 10; Length 118;

Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 113 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 172
Db 1 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 60
QY 173 TTTAAGCGCAACAAATTTATCGATAAATAAATCTTTTAGACGCAACAAATTTATAG 230
Db 61 TTTAAGCGCAACAAATTTATCGATAAATAAATCTTTTAGACGCAACAAATTTATAG 118

RESULT 4
US-09-895-435-7
; Sequence 7, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-7

Query Match 45.9%; Score 114.4; DB 10; Length 116;
Best Local Similarity 99.1%; Pred. No. 9.5e-14;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 113 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 172
Db 1 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 60
QY 173 TTTAAGCGCAACAAATTTATCGATAAATAAATCTTTTAGACGCAACAAATTTAT 228
Db 61 TTTAAGCGCAACAAATTTATCGATAAATAAATCTTTTAGACGCAACAAATTTAT 116

RESULT 5
US-09-895-435-8
; Sequence 8, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-8

Query Match 45.9%; Score 114.4; DB 10; Length 116;
Best Local Similarity 99.1%; Pred. No. 9.5e-14;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 113 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 172
Db 1 TAATACACACAAATATTATCTGTTGTTAGTTAGTTATTAATAATAATCTAGGG 60

Qy 173 TTTTAACGCAACACAATTTTATCGATAAATAAATACTTTTATAGCGCAACACAAATTTAT 228
|||
Db 61 TTTTAACGCAACACAATTTTATCGATAAATAAATACTTTTATAGCGCAACACAAATTTAT 116
|||

RESULT 6

```

US-09-895-435-9
Sequence 9, Application US/09895435
Publication No. US20030113837A1
GENERAL INFORMATION:
APPLICANT: Stewart, A. F.
APPLICANT: Zhang, Y.
APPLICANT: Hallett, B.
TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
FILE REFERENCE: 9882-012-999
CURRENT APPLICATION NUMBER: US/09/895,435
CURRENT FILING DATE: 2001-06-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 116
TYPE: DNA
ORGANISM: Bacillus thuringiensis
FEATURE:
NAME/KEY: N_region
LOCATION: 33..64
OTHER INFORMATION: n = a, t, g, or c
NAME/KEY: N_region
LOCATION: 81..100
OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-9

```

RESULT 7

```

US-09-895-435-10
; Sequence 10, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallset, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33 .. 74
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-10

Query Match          20.5%; Score 51; DB 10; Length 124;
Best Local Similarity 65.6%; Pred. No. 0.35;
Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps

```

Qy	113	TAATACACACAAATATTAATTTGTTGCTATTA	-----GGTGTATATAATAATA	162
Db	1	TAATACACACAAATATTAATTTGTTGCTATTA	AAAAAAAAAAAAAAAAAAAAAAAA	60
Qy	163	AATCTAGGGGTTTAAACGCAACACAAATTTATCGATAAATAAATACTTTTATAGACGCAACACA		222
Db	61	NNNNNNNNNNNNNACGCAACACAAATTTA-CGATAAATAAATACTTTTATAGACGCAACACA		119
Qy	223	ATTTA	227	
Db	120	ATTTA	124	

RESULT 8

```

US-10-424-599-137579
/ Sequence 137579, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa Thomas J
/ APPLICANT: Kovalic David K
/ APPLICANT: Zhou Yihua
/ APPLICANT: Cao Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 137579
/ LENGTH: 761
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_95242C.1
US-10-424-599-137579

```

RESULT 9

US-10-311-455-35/c
; Sequence 35, Application US/10311455
; Publication NO. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424

```
; SEQ ID NO 35
; LENGTH: 9905
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-35

Query Match          19.7%; Score 49; DB 14; Length 9905;
Best Local Similarity 51.1%; Pred. No. 3.2; Mismatches 0; Gaps 0;
Matches 115; Conservative 0;

QY 7 CGCGAGCATTTCCGGAAGAAACCCAGCTAAGAAATCAGAGTTAAAGAAATCAGAAATA 66
Db 9905 CCTAAACATCAACAAACAAATAATACCAATATTTTCCCGCTTACATCAAAATTA 9846

QY 67 TATCATTTCTTCGACACATACATGTTCTTTTATACAAAATTAATACACACAAT 126
Db 9845 TATCTTTTCATCTTTATACATACAAATTAATCCCATCTTAACACCTAAATATTA 9786

QY 127 ATTAATTTGTTGTTAGTGTCTTATAATAATAATAATCTAGGGTTTAAAGCAACACA 186
Db 9785 CATAAAAATAAATCTCTTAATTTATAAAAAATAAAAAATATCTTTTAAAAAAACAATA 9756

QY 187 ATTTATCGTAATAATAATACITTTTAGACGCAACACAAATTTATAGA 231
Db 9725 ACATAAAAAATAAACAATAATTAATACATAAACAACACTTAA 9661

RESULT 10
US-10-311-455-330/c
; Sequence 330, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 330
; LENGTH: 9642
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-330

Query Match          19.6%; Score 48.8; DB 14; Length 9642;
Best Local Similarity 51.4%; Pred. No. 3.5; Mismatches 0; Gaps 0;
Matches 113; Conservative 0;

QY 12 AGCATTTCCGGAAGAAACCCAGCTAAGAAATCAGAGTTAAAGAAATCAGAAATATATCA 71
Db 3816 AACATTTCCAAACAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3757

QY 72 TTATTCCTTCGACACATACATGTTCTTTTATACAAAATTAATACACACAATATTA 131
Db 3756 TACTTAAACCAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3697

QY 132 TTGTTGTTGTTAGTGTCTTATAATAATAATAATCTAGGGTTTAAAGCAACACAATTTA 191
Db 3696 ATATTAAAAAACCCACATATAAATAATATTTTAAAAATTTAAATCTACGAAATATA 3637

; SEQ ID NO 36
; LENGTH: 9905
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-36

Query Match          19.4%; Score 48.2; DB 14; Length 9905;
Best Local Similarity 53.4%; Pred. No. 4; Mismatches 88; Indels 0; Gaps 0;
Matches 101; Conservative 0;

QY 36 AAGAAATCAGAGTTAAAAAATCAGAAAAATATATCATTTCTTCGACACACATGTTTC 95
Db 2521 AAAATATTTAAATCTAAATTAATCTAAATTAATTAATTAATTAATTAATTAATTAAT 2462

QY 96 TTTTATATACAAAAATAATAACACAAATATTAATTTGTTGTTAGTGTCTTATTAAT 155
Db 2461 AATAAATATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2402

QY 156 AAATATAATCTAGGGTTTAAAGCAACACAATTTATCGATAAATAAATAAATAAATAA 215
Db 2401 AAAATATATATTTATTTATTTATATTAATAAATAAATAAATAAATAAATAAATAAATAA 2342

QY 216 CAACACAAT 224
Db 2341 AACATAAT 2333

RESULT 12
US-10-240-485-162/c
; Sequence 162, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
```

```

; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 162
; LENGTH: 18585
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-162

Query Match      19.0%; Score 47.2; DB 14; Length 18585;
Best Local Similarity 53.2%; Pred. No. 8.8;
Matches 100; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 42 ATCAGAGTAAATAATCAGAAATATATCATATTCCTTGACATACATGTTCTTTT 101
DB 8305 ATAAATTTATATAATAATTAATAATTAATAATTAATAATTAATAATTAATA 8246
QY 102 TATACAAAAATAATACACACAAATTAATTAATGTTGTTATAGGTGTTATAATAT 161
DB 8245 TAAATTTTAAAAAAAACATTTCAACCATATTTTATTAATTTTATAACAAAA 8186
QY 162 AAATCTAGGGTTTACGCAACACAAATTTATCGATAATAATATCTTTTACGCAACAC 221
DB 8185 ATATATAACTATTTAACATAATAATAATAATTTAAATTTAAATTTTAAACGAAAA 8126
QY 222 AATTTTATA 229
DB 8125 AAAATAAA 8118

RESULT 13
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match      19.0%; Score 47.2; DB 14; Length 3673778;
Best Local Similarity 54.7%; Pred. No. 42;
Matches 94; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 66 ATATCATTTTCTTGACATACATGTTCTTTTATACAAAAATAATATACACAA 125
DB 2632005 ATATTATGTTTATAAATATATATATTTTATTAATAAGAAATATTAATATA 2632064
QY 126 TATTAAATGTTGTTATAGGTGTTATATAATAATATCTAGGGTTTACGCAACAC 185
DB 2632065 ATAAATTTTATTATATATATAGTAAGTATATAAATTAATAATTTGAATAAAT 2632124
QY 186 AATTTATCGATAAATAATCTTTTACGCAACCAATTTATACACGCGGA 237
DB 2632125 TATGTAATAATAGATATATATTTTATTAATAAGATTTTATAGTTATGGA 2632176

RESULT 14
US-10-311-455-3/C
; Sequence 3, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 3
; LENGTH: 6072
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-3

Query Match      18.8%; Score 46.8; DB 14; Length 6072;
Best Local Similarity 52.6%; Pred. No. 7.6;
Matches 102; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 36 AAGAAATCAGAGTTTAAAAATCAGAAATATATATATTCCTTGACATACATGTTTC 95
DB 4174 AAAAAAATCTACCGAATAAATTCCTTAATCTTTTATTTACCCCACTATCATCTCTC 4115
QY 96 TTTTATTTATACAAAAATAATATACACAAATATTAATTCGTTGTTATAGGTGTTATAAT 155
DB 4114 TTCAATTTTACAAAAACAAATAAATAATTAATAATTTATTCGTTTCTACAAAT 4055
QY 156 AAATATAAATCTAGGGTTTAAACGCAACAAATTTATCGATAAATAATCTTTTAGACG 215
DB 4054 TAAATATAATCGAAATTTTTCCTCCCAACATAAAAAAATAAATACTAAATAAACT 3995
QY 216 CAACCAATTTATA 229
DB 3994 AAACATACATAATA 3981

RESULT 15
US-10-311-455-426/C
; Sequence 426, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Dete
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 426
; LENGTH: 6980
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-426
Query Match      18.8%; Score 46.8; DB 14; Length 6980;
Best Local Similarity 53.2%; Pred. No. 7.9;
Matches 99; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 16 TTTCGGAAAAAACCCACGCTAAGAAATCAGAGTTAAATAATCAGAAATATATCATTTAT 75
DB 6918 TTTCACAAACCAAAATTCAAATCTATCTTAAATTTAAATAAAATAACCCACACA 6859

QY 76 TCCTTGACACATACATGTTCTTTTATACAAAAATAATACACACAAATTAATTGT 135
DB 6858 TCTCTAAACATTTTATTTTATATTAAATAATACGACATAAATATATATAATTTT 6799

QY 136 GTTGATTAGGTGTTATTAATAATAATCTAGGGGTTTAAACGACACAAATTTATCGA 195
DB 6798 ATTTTATAAATCTAAAAAAAATTCCTTTTAATTAATTAATAAATAATATAAA 6739

QY 196 TAAATA 201
DB 6738 AAAATA 6733

```

Search completed: February 28, 2004, 09:12:59
Job time : 285.282 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:13:57 ; Search time 3916.83 Seconds
(without alignments)
2236.845 Million cell updates/sec

Title: US-09-895-435A-4

Perfect score: 249

Sequence: 1 ggggtaccgcagcattctg.....gacggagggaatcacatg 249

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:*

- 1: /cgn2_6/ptodata/2/pna/PCTUS COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US66 COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US080 COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US081 COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US082 COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US083 COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US084 COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US085 COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US086 COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US087 COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US088 COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US089 COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US090 COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091 COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US093A COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US093B COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US094 COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US095A COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US095B COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US095C COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US095D COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US096A COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US096B COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US096C COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US096D COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US096E COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US097A COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US097B COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US097C COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US098A COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US098B COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US098C COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US098D COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US099A COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US099B COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US099C COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US099D COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US099E COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US099F COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US100A COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US106 COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Result No.	Score	Query Match	Length	DB ID	Description
1	249	100.0	249	36	US-09-895-435A-4
2	249	100.0	249	36	US-09-895-435A-4
3	244	98.0	244	36	US-09-895-435A-1
4	244	98.0	244	36	US-09-895-435A-1
5	118	47.4	118	36	US-09-895-435A-2
6	118	47.4	118	36	US-09-895-435A-2
7	114.4	45.9	116	36	US-09-895-435A-7
8	114.4	45.9	116	36	US-09-895-435A-7
9	114.4	45.9	116	36	US-09-895-435A-8
10	114.4	45.9	116	36	US-09-895-435A-8
11	64	25.7	116	36	US-09-895-435A-9
12	64	25.7	116	36	US-09-895-435A-9
13	51.6	20.7	382	34	US-09-837-604A-6057
14	51.6	20.7	382	34	US-09-837-604B-6057
15	51.6	20.7	382	75	US-60-197-872-65927
16	51.6	20.7	484	37	US-09-912-293-117429
17	51	20.5	124	36	US-09-895-435A-10
18	51	20.5	124	36	US-09-895-435A-10
19	51	20.5	90859	22	US-09-534-859-183
20	51	20.5	90859	33	US-09-803-726-183
21	50.2	20.2	761	51	US-10-424-599-137579
22	50	20.1	528	35	US-09-865-439A-100898
23	50	20.1	528	76	US-60-207-458-144819
24	50	20.1	43545	1	PCT-US01-32202-134
25	50	20.1	43545	29	US-09-687-837-134
26	50	20.1	43545	53	US-10-663-538-134
27	49.8	20.0	33292	28	US-09-662-254A-27
28	49.8	20.0	33292	28	US-09-662-254A-27
29	49.8	20.0	33292	28	US-09-662-254B-27
30	49.8	20.0	33292	54	US-10-706-635-27
31	49.8	20.0	50000	28	US-09-662-254A-23
32	49.8	20.0	50000	28	US-09-662-254A-23
33	49.8	20.0	50000	28	US-09-662-254B-23
34	49.8	20.0	50000	54	US-10-706-635-23
35	49.6	19.9	474	18	US-09-304-517A-265446
36	49.6	19.9	474	19	US-09-371-146A-264381
37	49.6	19.9	474	24	US-09-565-240-12976
38	49.6	19.9	474	27	US-09-654-617-70590
39	49.6	19.9	474	29	US-09-684-016-70590
40	49.6	19.9	474	42	US-09-985-678-265446
41	49.6	19.9	474	70	US-60-141-134-3385
42	49.2	19.8	387	35	US-09-865-419A-40417
43	49.2	19.8	387	76	US-60-208-063-23473
44	49	19.7	281	18	US-09-304-517A-264727
45	49	19.7	281	19	US-09-371-146A-263669

Query Match 100.0%; Score 249; DB 36; Length 249;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCACGCTAAGAAATCAGAGTTAAAAAATCAG 60
DB 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCACGCTAAGAAATCAGAGTTAAAAAATCAG 60
QY 61 AAAATATATCATTTATTCCTTCGACACATACATGTTCTTTTATACAAAAAATAATACAA 120
DB 61 AAAATATATCATTTATTCCTTCGACACATACATGTTCTTTTATACAAAAAATAATACAA 120
QY 121 CACAATATTAATGTTGTTGTTATAGGTTTATAATAATAAATCTAGGGGTTTAAACGC 180
DB 121 CACAATATTAATGTTGTTGTTATAGGTTTATAATAATAAATCTAGGGGTTTAAACGC 180
QY 181 AACACAATTTATCGATAAATAAATACTTTTATAGCGCACACAAATTTATAGCGCGGAGGA 240
DB 181 AACACAATTTATCGATAAATAAATACTTTTATAGCGCACACAAATTTATAGCGCGGAGGA 240
QY 241 AATCACATG 249
DB 241 AATCACATG 249

RESULT 2
US-09-895-435A-4
; Sequence 4, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435A-4

Query Match 100.0%; Score 249; DB 36; Length 249;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCACGCTAAGAAATCAGAGTTAAAAAATCAG 60
DB 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCACGCTAAGAAATCAGAGTTAAAAAATCAG 60
QY 61 AAAATATATCATTTATTCCTTCGACACATACATGTTCTTTTATACAAAAAATAATACAA 120
DB 61 AAAATATATCATTTATTCCTTCGACACATACATGTTCTTTTATACAAAAAATAATACAA 120
QY 121 CACAATATTAATGTTGTTGTTATAGGTTTATAATAATAAATCTAGGGGTTTAAACGC 180
DB 121 CACAATATTAATGTTGTTGTTATAGGTTTATAATAATAAATCTAGGGGTTTAAACGC 180
QY 181 AACACAATTTATCGATAAATAAATACTTTTATAGCGCACACAAATTTATAGCGCGGAGGA 240
DB 181 AACACAATTTATCGATAAATAAATACTTTTATAGCGCACACAAATTTATAGCGCGGAGGA 240
QY 241 AATCACATG 249
DB 241 AATCACATG 249

RESULT 3
US-09-895-435-1
; Sequence 1, Application US/09895435
; GENERAL INFORMATION:

ALIGNMENTS

RESULT 1
US-09-895-435-4
; Sequence 4, Application US/09895435
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-4

```
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 244
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-1

Query Match
Best Local Similarity 98.0%; Score 244; DB 36; Length 244;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGTACGCCAGCATTTCCGAAAAAACCACGCTAAGAAAAATCAGAGTTAAAAAATCAGAA 62
DB 1 GGTACGCCAGCATTTCCGAAAAAACCACGCTAAGAAAAATCAGAGTTAAAAAATCAGAA 60
QY 63 AATATATCATATTCCTTGACACATACATGTTCTTTTTTATACAAAAAATAATACACA 122
DB 61 AATATATCATATTCCTTGACACATACATGTTCTTTTTTATACAAAAAATAATACACA 120
QY 123 CAATATTAAATGTTGTTAGTGTATTAATAATAATAATTAATCTAGGGGTTTAAACGCA 182
DB 121 CAATATTAAATGTTGTTAGTGTATTAATAATAATAATTAATCTAGGGGTTTAAACGCA 180
QY 183 CACAATTTTCATATAAATAATCTTTTAGCGCAACACATTTTATAGCGGGGAA 242
DB 181 CACAATTTTCATATAAATAATCTTTTAGCGCAACACATTTTATAGCGGGGAA 240
QY 243 TCAC 246
DB 241 TCAC 244

RESULT 4
US-09-895-435A-1
; Sequence 1, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 244
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435A-1

Query Match
Best Local Similarity 98.0%; Score 244; DB 36; Length 244;
Matches 244; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GGTACGCCAGCATTTCCGAAAAAACCACGCTAAGAAAAATCAGAGTTAAAAAATCAGAA 62
DB 1 GGTACGCCAGCATTTCCGAAAAAACCACGCTAAGAAAAATCAGAGTTAAAAAATCAGAA 60
QY 63 AATATATCATATTCCTTGACACATACATGTTCTTTTTTATACAAAAAATAATACACA 122
DB 61 AATATATCATATTCCTTGACACATACATGTTCTTTTTTATACAAAAAATAATACACA 120
QY 123 CAATATTAAATGTTGTTAGTGTATTAATAATAATAATTAATCTAGGGGTTTAAACGCA 182
```

```
DB 121 CAATATTAAATGTTGTTAGTGTATTAATAATAATAATAATCTAGGGGTTTAAACGCA 180
QY 183 CACAATTTTCGATAAATAAATACTTTTAGCGCAACACAAATTTATAGACGGGAGGAAA 242
DB 181 CACAATTTTCGATAAATAAATACTTTTAGCGCAACACAAATTTATAGACGGGAGGAAA 240
QY 243 TCAC 246
DB 241 TCAC 244

RESULT 5
US-09-895-435-2
; Sequence 2, Application US/09895435
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-2

Query Match
Best Local Similarity 47.4%; Score 118; DB 36; Length 118;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 TAATACACACAAATATTGTTGTTAGTGTATTAATAATAATAATAATCTAGGGG 172
DB 1 TAATACACACAAATATTGTTGTTAGTGTATTAATAATAATAATAATCTAGGGG 60
QY 173 TTTTACGCAACACAAATTTATCGATAAATAATACTTTTAGCGCAACACAAATTTATAG 230
DB 61 TTTTACGCAACACAAATTTATCGATAAATAATACTTTTAGCGCAACACAAATTTATAG 118

RESULT 6
US-09-895-435A-2
; Sequence 2, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 118
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435A-2

Query Match
Best Local Similarity 47.4%; Score 118; DB 36; Length 118;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 TAATACACACAAATATTGTTGTTAGTGTATTAATAATAATAATAATCTAGGGG 172
DB 1 TAATACACACAAATATTGTTGTTAGTGTATTAATAATAATAATAATCTAGGGG 60
QY 173 TTTTACGCAACACAAATTTATCGATAAATAATACTTTTAGCGCAACACAAATTTATAG 230
DB 61 TTTTACGCAACACAAATTTATCGATAAATAATACTTTTAGCGCAACACAAATTTATAG 118
```

7

```
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; LOCATION: 33..64
; OTHER INFORMATION: n = a, t, g, or c
; NAME/KEY: N region
; LOCATION: 81..100
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435a-9

Query Match      25.7%; Score 64; DB 36; Length 116;
Best Local Similarity 55.2%; Pred. No. 0.038;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 113 TAATACACACAAATTAATTCGTAATAGGTTATATAATAATAATCTAGGG 172
Db 1 TAATACACACAAATTAATTCGTTATATANNNNNNNNNNNNNNNNNNNNNNNN 60

Qy 173 TTAAACGCAACAATTTATCGATAAATACTTTTAGCGCAACAATTTAT 228
Db 61 NNNNACGCAACAATTTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 116

RESULT 12
US-09-895-435a-9
; Sequence 9, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; PRIOR FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33..64
; OTHER INFORMATION: n = a, t, g, or c
; FEATURE:
; NAME/KEY: N region
; LOCATION: 81..100
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435a-9

Query Match      25.7%; Score 64; DB 36; Length 116;
Best Local Similarity 55.2%; Pred. No. 0.038;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 113 TAATACACACAAATTAATTCGTTATAGGTTATATAATAATAATACTAGGG 172
Db 1 TAATACACACAAATTAATTCGTTATANNNNNNNNNNNNNNNNNNNNNNNNNN 60

Qy 173 TTAAACGCAACAATTTATCGATAAATACTTTTAGCGCAACAACAATTTAT 228
Db 61 NNNNACGCAACAATTTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 116

RESULT 13
US-09-837-604A-69057
; Sequence 69057, Application US/09837604A
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: De La Pena, Robert C.
; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604A
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 69057
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(382)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-048-Q6-N1-G9
US-09-837-604A-69057

Query Match      20.7%; Score 51.6; DB 34; Length 382;
Best Local Similarity 51.5%; Pred. No. 7.6;
Matches 102; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 15 ATTTCGGAAGAAACCAACGCTAGAAATCAGAGTTTAAATCAGAAATATATATATCA 74
Db 184 ATTTTAAAGAAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 243

Qy 75 TTCTTGACACATACATGTTCTTTTATACAAAAATAATACACACAAATTAATTG 134
Db 244 AAAAAATTTTAAAGGAATAAAATTTTAAATTTTAAATTAAGGGAGTTAAAAAAGTT 303

Qy 135 TGTTGATTAGGTGTTATATAATAATACTAGGGTTTAAAGCAACAATTTATCG 194
Db 304 TTTTNTNNGAAAAAAGGGAGGNTTTTATATAAAATTTTATATAAAATTTT 363

Qy 195 ATAAATAATACTTTTAG 212
Db 364 AAAAAAATAATTTTAG 381

RESULT 14
US-09-837-604B-69057
; Sequence 69057, Application US/09837604B
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: De La Pena, Robert C.
; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604B
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 69057
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(382)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3598-048-Q6-N1-G9
US-09-837-604B-69057

Query Match      20.7%; Score 51.6; DB 34; Length 382;
Best Local Similarity 51.5%; Pred. No. 7.6;
Matches 102; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:22:37 ; Search time 42.7439 Seconds
(without alignments)
981.354 Million cell updates/sec

Title: US-09-895-435A-4

Perfect score: 249

Sequence: 1 ggggtaccgcagcatttcg.....gacgcggaggaaatcacatg 249

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 164199 seqs, 84230614 residues

Total number of hits satisfying chosen parameters: 328398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New.*

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	17.3	641	6	US-10-767-701-20322
2	41.8	16.8	380963	6	Sequence 20322, A
3	41.6	16.7	350570	6	Sequence 5, Appli
4	40	16.1	803755	6	Sequence 146, App
5	39.6	15.9	43868	6	Sequence 10700, A
6	39.4	15.8	103977	6	Sequence 10680, A
7	39.2	15.7	52896	6	Sequence 10722, A
8	39.2	15.7	99445	6	US-10-767-471-10810
9	38.2	15.3	166020	6	Sequence 10810, A
10	38.2	15.3	1730242	6	Sequence 10806, A
11	38	15.3	41514	6	Sequence 10805, A
12	38	15.3	1730242	6	Sequence 10654, A
13	37.8	15.2	447	6	Sequence 10805, A
14	37.6	15.1	36106	6	Sequence 17775, A
15	37.6	15.1	156652	6	Sequence 10713, A
16	37	14.9	99971	6	Sequence 48, Appli
17	36.6	14.7	13621	6	Sequence 10561, A
18	36.6	14.7	92834	6	US-10-765-790-94
19	36.6	14.7	189817	6	Sequence 10676, A
20	36.4	14.6	15324	6	Sequence 10675, A
21	36.4	14.6	104249	6	Sequence 10559, A
22	36.2	14.5	40000	6	US-10-767-471-10652
23	36.2	14.5	47459	6	Sequence 10915, A
24	36	14.5	398800	1	Sequence 10701, A
25	35.8	14.4	38719	6	US-10-767-471-10701
26	35.8	14.4	149591	6	PCT-US03-32805-1
					Sequence 43, Appli
					Sequence 10601, A

ALIGNMENTS

RESULT 1

US-10-767-701-20322

; Sequence 20322, Application US/10767701

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 20322

; LENGTH: 641

; TYPE: DNA

; ORGANISM: Sorghum bicolor

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB5048-003-R1-XP2-D10

US-10-767-701-20322

Query Match 17.3%; Score 43; DB 6; Length 641;

Best Local Similarity 54.0%; Pred. No. 1;

Matches 88; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 65 TATATCATTTTCCTTGACACATACATGTTCTTTTATACAAAAATAATACACACA 124

Db 118 TGTATCATGTTTAAAGAAATPAAAAATATTTTATCTCGAAAAAGAGATCTAGA 177

QY 125 ATATTAATGTTGTTTATAGTGTATATAATAATATAATCTAGGGTTTAAACCAACA 184

Db 178 AGATTTTAAGTTTATGTTTACATTATAATAAAAAAATTTCTTGAAGTGAATACCTCC 237

QY 185 CAATTTCGTAATAATAATCTTTTAGCGCACACAAATTTA 227

Db 238 AATCTTATGATCTAGTTATTTTAGAAGAAATTAATTTA 280

RESULT 2

US-10-765-790-5/c

; Sequence 5, Application US/10765790

; GENERAL INFORMATION:

; APPLICANT: Bayer Healthcare LLC

; APPLICANT: Beard, Chris

; APPLICANT: Burgess, Chris

; APPLICANT: Cannon, Allison

; APPLICANT: Harvey, Jeanne

; APPLICANT: Lechner, John F.

; APPLICANT: Li, Zheng

QY 64 ATATATCAATTCTTGGACACATACATGTTCTTTTATACAAAAAATAACACAC 123
DB 18673 GTAAGATAGGTATTTAAATCAGTGAIGTACTTGAATGAACATAAAATAAGTAAC 18732
QY 124 AATATTAATTGTTGTTAGTGTATTAATAATAATAATCTAGGGTTTAA 177
DB 18733 AAAATTAACCAAGACCTTTAAATATATATCTGACAATCACTTTGTAGGTAGAA 18786

RESULT 6
US-10-767-471-10722
; Sequence 10722, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10722
; LENGTH: 109977
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(109977)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10722

Query Match 15.8%; Score 39.4; DB 6; Length 109977;
Best Local Similarity 56.6%; Pred. No. 5.1;
Matches 73; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 53 AAAATCAGAAAAATATATCATTCCTTCACACATACATGTTCTTTTATACAAAAA 112
DB 33890 AAAATGTAAATATAGCAATATTATAGAACATAAGCTTAAATATGTTTATATAAA 33949
QY 113 TAATACACACAAATTAATTGTTGTTATAGGTGTTATATAATAATAATCTAGGG 172
DB 33950 GTAAATACAAACTATATCTGTGGAAATGTGTGTATATATATATATATGCGAG 34009
QY 173 TTTAACGCA 181
DB 34010 TCAGAGGTA 34018

RESULT 7
US-10-767-471-10810
; Sequence 10810, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10810
; LENGTH: 52896
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10810

Query Match 15.7%; Score 39.2; DB 6; Length 52896;
Best Local Similarity 56.1%; Pred. No. 5.5;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 74 ATTCCCTGCACATACATGTTCTTTTATACAAAAAATAACACAAATTAAT 133

DB 6920 ATGACTAGACATTACTTTTGTGTTTTTATATATATATATAAAACAATAATATCTATT 6979
QY 134 GTGTTGTTATAGGTGTTATAATAATAATAAAATCTAGGGTTTAAACGCAACAATTTATC 193
DB 6980 ATATATATTTATATAAACAAATAATATATTATTATATATATTATATAAAACAATAATA 7039
QY 194 GATAAATAATA 205
DB 7040 TATATATTTATA 7051

RESULT 8
US-10-767-471-10714
; Sequence 10714, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10714
; LENGTH: 99445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10714

Query Match 15.7%; Score 39.2; DB 6; Length 99445;
Best Local Similarity 56.1%; Pred. No. 5.6;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 74 ATTCCCTGCACATACATGTTCTTTTATACAAAAAATAACACAAATTAAT 133
DB 95498 ATGACTAGACATTACTTTTGTGTTTTTATATATATATAAAACAATAATATCTATT 95557
QY 134 GTGTTGTTATAGGTGTTATAATAATAATAATCTAGGGTTTAAACGCAACAATTTATC 193
DB 95558 ATATATTTATATAAACAAATAATATATTATTATATATATAAAACAATAATA 95617
QY 194 GATAAATAATA 205
DB 95618 TATATATTTATA 95629

RESULT 9
US-10-767-471-10806/c
; Sequence 10806, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10806
; LENGTH: 166020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(166020)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables
US-10-767-471-10806

Query Match 15.3%; Score 38.2; DB 6; Length 166020;
Best Local Similarity 49.7%; Pred. No. 8.7;
Matches 97; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 37 AGAAATCAGAGTTAAATAATCAGAAAAATATATCAATTATTCTTGCACATACATGTTCT 96

FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 53128
SEQ ID NO 17775
LENGTH: 447
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: LIB3476-047-P1-K1-D11
US-10-767-701-17775

Query Match 15.2%; Score 37.8; DB 6; Length 447;
Best Local Similarity 48.8%; Pred. No. 9.5;
Matches 102; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 23 AAAAAACACGCTAAGAAATACAGAGTTAAATAATCAGAAATATATATATTCCTTGA 82
DB AAAAACTTAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAA 369
QY 83 CACATACATGTTCTTTTATACAAAAATAATACACACAAATTAATTTGTTGTAT 142
DB AAAAAAATATATATAAAAAAACATTAAATAATTTTAAAAATAACTAAACAAATAA 309
QY 143 TAGGTGTTATATATATATATCTAGGGTTTAAACGCAACACAAATTTATCGATAATAA 202
DB AATATAAAAAATAAT 249
QY 203 ATACTTTTAGCGCAACAAATTTATAGA 231
DB ATAGAAAAAATAGAACACTATATAGACA 220

RESULT 14

US-10-767-471-10713/c
Sequence 10713, Application US/10767471
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001505
CURRENT APPLICATION NUMBER: US/10/767,471
CURRENT FILING DATE: 2004-01-30
NUMBER OF SEQ ID NOS: 50231
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10713
LENGTH: 36106
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(36106)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10713

Query Match 15.1%; Score 37.6; DB 6; Length 36106;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 90; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 44 CAGAGTTAAAAATCAGAAAAATATATATCTTCTTGACACATACATGTTCTTTTGA 103
DB 30780 CTGTTTCAATGATCAAGTATATATATATATATATATATATATATATATATATAT 30721
QY 104 TACAAAAATAATACACAAATATTAATGTTGTTAGGTATTAATAATAATAA 163
DB 30720 TAAAAAGTAT-ATATATGATATATATATCTTTTGTATATAAAGTATATATATAT 30662
QY 184 ATCTAGGGGTTTACGCAACCAATTTATCGATAATAA 203
DB 30661 ATATACATTTTTTTTGTATATAAAGTATACAAAAA 30622

RESULT 15

US-10-765-790-48
Sequence 48, Application US/10765790
GENERAL INFORMATION:
APPLICANT: Bayer Healthcare LLC
APPLICANT: Beard, Chris
APPLICANT: Burgess, Chris
APPLICANT: Gannon, Allison
APPLICANT: Harvey, Jeanne
APPLICANT: Lechner, John F.
APPLICANT: Li, Zheng
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
FILE REFERENCE: 1657/2035
CURRENT APPLICATION NUMBER: US/10/765,790
CURRENT FILING DATE: 2004-01-27
PRIOR APPLICATION NUMBER: US 10/737,082
PRIOR FILING DATE: 2003-12-16
NUMBER OF SEQ ID NOS: 300
SOFTWARE: PatentIn version 3.2
SEQ ID NO 48
LENGTH: 156652
TYPE: DNA
ORGANISM: Homo sapiens
US-10-765-790-48

Query Match 15.1%; Score 37.6; DB 6; Length 156652;
Best Local Similarity 53.4%; Pred. No. 11;
Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 101 TTATACAAAAATAATACAAACAATATTAATTTGTTGTTATAGGTGTTATAATAATA 160
DB 57943 TTATCTATAAATATGAAACATCAAACTATATATGTTTCATTTATTTGATTATAATA 58002
QY 161 TAAATCTAGGGTTTAAACGCAACAAATTTATCGATAATAATAATACTTTTAGACGAACA 220
DB 58003 TCTTTATGAGATATGAATAAATGTTTATGTTCTCTATTAAAGATATGATGATGA 58062
QY 221 CAATTTATAGACGCGAGGAAATCAT 248
DB 58063 AAGTTTATATCTTTGGAAGATTACAT 58090

Search completed: February 28, 2004, 10:50:50
Job time : 51.7439 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:05:57 : Search time 2938.47 Seconds
(without alignments)
2530.460 Million cell updates/sec

Title: US-09-895-435A-4

Perfect score: 249

Sequence: 1 ggggtaccgcagcatttcg.....gagcggagggaatcacatg 249

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_tod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	58.6	23.5	1179	13	BX417261
2	57.4	23.1	1101	29	CNS0022U
3	56.6	22.7	926	29	CNS006YV
4	55.2	22.2	1165	13	BX338369

5	55	22.1	987	29	CNS014PQ
6	54.8	22.0	924	29	CNS0206S
7	53.6	21.5	989	13	BX33264
8	53.2	21.4	1072	13	BX359378
9	53	21.3	976	29	CNS009G1
C 10	53	21.3	1201	13	BX334085
C 11	52.8	21.2	734	29	CNS010MP
C 12	52.8	21.2	1201	13	BX462207
C 13	52.6	21.1	832	29	CNS0110K
C 14	52.6	21.1	1201	13	BX458623
15	52.4	21.0	240	9	AU074022
16	52.4	21.0	1201	29	CNS01038
17	52.2	21.0	998	13	BX436885
18	51.8	20.8	1092	13	BX359398
C 19	51.4	20.6	1101	29	CNS000B8
C 20	51	20.5	1001	29	CNS0155H
C 21	51	20.5	1021	29	CNS00133
C 22	50.8	20.4	840	29	CNS010TW
C 23	50.8	20.4	1203	29	CNS015WU
24	50.6	20.3	867	29	CNS074W9
25	50.6	20.3	1201	13	BX439779
C 26	50.4	20.2	637	29	CNS036CC
27	50.4	20.2	1101	29	CNS003BD
28	50.2	20.2	318	28	AQ092950
29	50.2	20.2	420	12	BM891761
C 30	50.2	20.2	577	29	AQ393140
C 31	50.2	20.2	973	29	CNS008OF
32	50.2	20.2	1201	13	BX401903
33	50.2	20.2	1201	13	BX458623
34	50	20.1	427	13	BX403499
C 35	50	20.1	639	29	CE144836
36	50	20.1	999	13	BX380865
37	50	20.1	1049	13	BX353130
38	50	20.1	1164	13	BX425839
39	49.8	20.0	1101	29	CNS017KE
40	49.6	19.9	1039	29	CNS006H7
C 41	49.6	19.9	1092	29	CNS020K7
42	49.4	19.8	976	14	CA972153
43	49.4	19.8	1179	13	BX417261
C 44	49.4	19.8	1311	10	BE420654
45	49.4	19.8	2039	13	BX447434

ALIGNMENTS

RESULT 1
BX417261/c
LOCUS BX417261 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE006YP16
DEFINITION S-PRIME, mRNA sequence.
ACCESSION BX417261
VERSION BX417261.1 GI:30650375
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1179)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6333.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE006DH08QP1&cluster=6333.r. Contact :
feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DE006DH08QP1.

```
FEATURES
  source
    Location/Qualifiers
      1..1179
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DE006YP16"
        /tissue_type="PLACENTA"
        /clone_lib="Homo sapiens PLACENTA"
        /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
  Query Match      23.5%; Score 58.6; DB 13; Length 1179;
  Best Local Similarity 48.2%; Pred. No. 0.34;
  Matches 94; Conservative 21; Mismatches 80; Indels 0; Gaps 0;

QY 16 TTTCGAAAAAACCACGCTAAGAAATCAAGTTAAAAATCAGAAAAATATATCATAT 75
Db 1150 TTTCGAAAAAAGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATTT 1101

QY 76 TCCITGACACATACATGCTCTTTTATACAAAAATAATACACACATATTAATTGT 135
Db 1100 TAWTTTTTTTWWTTTTTTTWTWTWTWTAAATAATTAATWAAATAATWAAATTT 1041

QY 136 GTGTATTAGTGTATTATATAAATAAATCTAGGGTTTAAACGACACAAATTTATCGA 195
Db 1040 TTTTWTATTTTTTTTTTTTAAAWAAAWAATTTATTTTAAAWWAAAWATTTTGTAT 981

QY 196 TAAATAAATACTTTT 210
Db 980 AAAAAATTTTWTWT 966

RESULT 2
CNS0022U
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence SP6 end of BAC
  BACN01007 of DrosBAC library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL097152
VERSION
  AL097152.1 GI:5608763
KEYWORDS
  Drosophila melanogaster (fruit fly)
SOURCE
  Drosophila melanogaster
  ORGANISM
    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
    Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
    Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 1101)
  Genoscope.
  Direct Submission
  Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (EDGP) -
  http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
  library (DrosBAC) was made by Alain Billaud at CSEH (Centre
  d'Etude du Polymorphisme Humain) with funding provided by a WRC
  project grant. The DNA was prepared from embryos by Alain Bucheton
  and Genevieve Payan. It has been constructed in the vector
  pBelOBAC11.
FEATURES
  source
    Location/Qualifiers
      1..1101
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACN01007"
        /clone_lib="DrosBAC"
        /plasmid="pBelOBAC11"

CNS0022U
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence SP6 end of BAC
  BACN01007 of DrosBAC library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL097152
VERSION
  AL097152.1 GI:5608763
KEYWORDS
  Drosophila melanogaster (fruit fly)
SOURCE
  Drosophila melanogaster
  ORGANISM
    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
    Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
    Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 1101)
  Genoscope.
  Direct Submission
  Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the European Drosophila Genome Project (EDGP) -
  http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
  library (DrosBAC) was made by Alain Billaud at CSEH (Centre
  d'Etude du Polymorphisme Humain) with funding provided by a WRC
  project grant. The DNA was prepared from embryos by Alain Bucheton
  and Genevieve Payan. It has been constructed in the vector
  pBelOBAC11.
FEATURES
  source
    Location/Qualifiers
      1..1101
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACN01007"
        /clone_lib="DrosBAC"
        /plasmid="pBelOBAC11"

CNS006YV
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence T7 end of BAC #
  BACR14P07 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL066199
VERSION
  AL066199.1 GI:4945066
KEYWORDS
  Drosophila melanogaster (fruit fly)
SOURCE
  Drosophila melanogaster
  ORGANISM
    Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
    Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
    Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 926)
  Genoscope.
  Direct Submission
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazuo Oosawa and
  Aaron Mammoss in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain Y; cn bw sp, the same strain used for the BDGP's
  Pi and EST libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
  source
    Location/Qualifiers
      1..926
        /organism="Drosophila melanogaster"
        /mol_type="genomic DNA"
        /db_xref="taxon:7227"
        /clone="BACR14P07"
        /clone_lib="RPCI-98"
        /note="end : T7"

ORIGIN
  Query Match      22.7%; Score 56.6; DB 29; Length 926;
  Best Local Similarity 43.3%; Pred. No. 0.86;
  Matches 90; Conservative 30; Mismatches 88; Indels 0; Gaps 0;
```

```
QY 22 AAAAAAACCGCTAGAGAAATCAGAGTTAAATAATCAGAAATATATATATTCCTTG 81
|||||
Db 489 AAAAAACAAACCCAMAWAAWAAATTTGWTAAATTTTAAATATATATATTTTAAA 548
|||||
QY 82 ACACATACATGTTCTTTTATATACAAAAATATATACAAATATATATTCGTGTA 141
:|||||
Db 549 WRATATATGTTGTTTATATATATATATATATATATATATATATATATATTTT 608
|||||
QY 142 TTAGTGTTTATATATATATATATATATATATATATATATATATATATATAT 201
|||||
Db 609 TCWGTATTTATATATATATATATATATATATATATATATATATATATAT 668
|||||
QY 202 AATACTTTTACAGCAACACAAATTTATA 229
:|||||
Db 569 WAWATAGRTDTGWAKRAATRAAW 696
|||||

RESULT 4
BX338369
LOCUS
DEFINITION BX338369 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI058YF14 5-PRIME, mRNA sequence.
ACCESSION BX338369
VERSION BX338369.1 GI:30343682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1165)
Li W.B., Gruber C., Jesses, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI058DC070F1.
Location/Qualifiers
1..1165
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI058YF14"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 22.2%; Score 55.2; DB 13; Length 1165;
Best Local Similarity 41.1%; Pred. No. 1.3;
Matches 78; Conservative 37; Mismatches 75; Indels 0; Gaps 0;

QY 22 AAAAAAACCGCTAGAGAAATCAGAGTTAAATAATCAGAAATATATATATTCCTTG 81
|||||
Db 815 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 874
|||||
QY 82 ACACATACATGTTCTTTTATATACAAAAATATATACAAATATATATTCGTGTA 141
:|||||
Db 875 WATATATATATATATATATATATATATATATATATATATATATATATATAT 934
|||||
QY 142 TTAGTGTTTATATATATATATATATATATATATATATATATATATATATAT 201
:|||||
Db 935 TTTTATATATATATATATATATATATATATATATATATATATATATATATAT 994
|||||
QY 202 AATACTTTTATA 211
:|||||
```

```
Db 995 TWATTTTTT 1004

RESULT 5
CNS014PQ
LOCUS
DEFINITION CNS014PQ 987 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL104456
VERSION AL104456.1 GI:5616067
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
library (Dros-BAC) was made by Alain Billaud at CRPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
Location/Qualifiers
1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN12P22"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

ORIGIN
Query Match 22.1%; Score 55; DB 29; Length 987;
Best Local Similarity 35.7%; Pred. No. 1.6;
Matches 75; Conservative 51; Mismatches 84; Indels 0; Gaps 0;

QY 22 AAAAAAACCGCTAGAGAAATCAGAGTTAAATAATCAGAAATATATATATTCCTTG 81
:|||||
Db 730 WAAAAAATAATTTAAATTTTAAAGAAATTTTAAATTTTAAATTTTAAATTTTAA 789
|||||
QY 82 ACACATACATGTTCTTTTATATACAAAAATATATACAAATATATATTCGTGTA 141
:|||||
Db 790 ATTAATATATTTTATATATATATATATATATATATATATATATATATATAT 849
|||||
QY 142 TTAGTGTTTATATATATATATATATATATATATATATATATATATATATAT 201
:|||||
Db 850 TAAATTTTWWAAATAATAGRTTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 909
|||||
QY 202 AATACTTTTACGCAACACAAATTTTATA 231
:|||||
Db 910 WWWAAATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 939
|||||

RESULT 6
CNS020Q6S
LOCUS
DEFINITION CNS020Q6S 924 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
157A04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL208909
VERSION AL208909.1 GI:7867728
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
```

ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE 1
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Bizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE 20359837
PUBMED 10899143

REFERENCE 3 (bases 1 to 924)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.

FEATURES
source
Location/Qualifiers
1..924
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="157A04"
/clone_lib="G"
/note="Genoscope sequence ID : COAG157BA028P1-end :
PUC-ori"

ORIGIN
Query Match 22.0%; Score 54.8; DB 29; Length 924;
Best Local Similarity 47.3%; Pred. No. 1.8;
Matches 86; Conservative 21; Mismatches 75; Indels 0; Gaps 0;
QY 24 AAAAAACCGCTAAGAAAATCAGAGTTAAAAAATCAGAAAATATATCATTTCTCTGAC 83
Db 693 AWAAAAATTWAAATWAT 752
QY 84 ACATACATGCTCTCTTTTATACAAAAATATATACACACAAATATATATGTTGTTATT 143
Db 753 AAAAAATTWTTTWTWAAAAAANWAAWATATATATATATATATATATATATATATAT 812
QY 144 AGGTGTATATAATA 203
Db 813 AWAAAWATA 872
QY 204 TA 205
Db 873 AA 874

RESULT 7
BX339264
LOCUS BX339264 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION Full-length cDNA libraries and normalization
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
ACCESSION BX339264
VERSION BX339264.1 GI:30337783
KEYWORDS EST.

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 989)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secre@genoscope.cns.fr; Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8045.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1070B08QPl&cluster=8045.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1070B08QPl.
Location/Qualifiers
1..989
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1070G16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 21.5%; Score 53.6; DB 13; Length 989;
Best Local Similarity 45.0%; Pred. No. 2.7;
Matches 98; Conservative 23; Mismatches 97; Indels 0; Gaps 0;
QY 19 CGGAAAAAACCCAGCTAAGAAATCAGAGTTAAAAAATCAGAAAATATATCATTTATCC 78
Db 626 CGGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 685
QY 79 TTGACACATACATGTTCTTTTATACAAAAAATAATACACACAAATATTAATTTGTT 138
Db 686 TWTWTTTWTTTTWTTTTWTTTTWTTTTWTTTTWTTTTWTTTTWTTTTWTTTTWTTTT 745
QY 139 GTATTAGTGTATATATATATATATATATCTAGGGTTTAAACGACACAAATATCGATAA 198
Db 746 SAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 805
QY 199 ATAAATACTTTTAGACGCAACAAATTTATAGACGCGG 236
Db 806 AAARGGGTYTTTWDGGGRTTKTTTAAARRRGG 843

RESULT 8
BX359378
LOCUS BX359378 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1056VP18 5-PRIME, mRNA sequence.
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
ACCESSION BX359378
VERSION BX359378.1 GI:30378326
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1072)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CSODI056DH09QPl.

FEATURES

source

1. 1072
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI056VP18"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 21.4%; Score 53.2; DB 13; Length 1072;
 Best Local Similarity 47.9%; Pred. No. 3.1;
 Matches 91; Conservative 18; Mismatches 81; Indels 0; Gaps 0;

QY 22 AAAAAACACGCTAGAAATCAGAGTTAAATTCAGAAATCAGAAATATATATATTCCTTG 81
 DB 531 AA 590

QY 82 ACACATACATGTTCTTTTATACAAAAATATACACAAATATTAATTTGTTGTA 141
 DB 591 AAAAAAAAAATTTTTTTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 650

QY 142 TTAGGTGTTAATAAATAATAAATCAGGGTTTAAACGCAACACAAATTTATCGATAAATA 201
 DB 651 AAAAAATTTTTTTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 710

QY 202 AATACCTTTA 211
 DB 711 TTTATTTT 720

RESULT 9
 CNS009G1
 LOCUS
 DEFINITION
 Drosophila melanogaster genome survey sequence TET3 end of BAC #
 BACR19J14 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL053529
 AL053529.1 GI:4935018
 GSS.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 876)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see http://www.fruitfly.org The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoshi Osoegawa and
 Aaron Mammosser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see http://www.fruitfly.org The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoshi Osoegawa and
 Aaron Mammosser in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 876
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR19J14"
 /clone_lib="RPCI-98"
 /note="end : TET3"

ORIGIN

Query Match 21.3%; Score 53; DB 29; Length 876;
 Best Local Similarity 45.9%; Pred. No. 3.7;
 Matches 95; Conservative 19; Mismatches 93; Indels 0; Gaps 0;

QY 22 AAAAAACACGCTAGAAATCAGAGTTAAATTCAGAAATCAGAAATATATATATTCCTTG 81
 DB 524 AATAAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 583

QY 82 ACACATACATGTTCTTTTATACAAAAATATACACAAATATTAATTTGTTGTA 141
 DB 584 AWTAAWTTAATTTATATAAATTTATATATAAATAAATAAATAAATAAATAAATA 643

QY 142 TTAGGTGTTAATAAATAATAAATCAGGGTTTAAACGCAACACAAATTTATCGATAAATA 201
 DB 644 TAAWATATATATATATTTTAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 703

QY 202 AATACCTTTTAGCGCAACAAATTTAT 228
 DB 704 TATATATTTTATATAAATAATATAW 730

RESULT 10

BX334085/c

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Reference

Authors

Title

Journal

Comment

BX334085 1201 bp mRNA linear EST 02-MAY-2003
 BX334085 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 CDNA clone CSODD006YL06 5-PRIME, mRNA sequence.

BX334085 BX334085.1 GI:30333283

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 6562.r For

more information about this cluster, see

<http://www.genoscope.cns.fr/>

cgi-bin/cluster.cgi?seq=CSODD006DF03QPl&cluster=6562.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CSODD006DF03QPl.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODD006YL06"

/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

LOCUS CDS110K 832 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BAC06G12 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL100526
VERSION AL100526.1 GI:5612137
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 832)
AUTHORS Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CRPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source Location/Qualifiers
1..832
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN06G12"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"
ORIGIN
Query Match 21.1%; Score 52.6; DB 29; Length 832;
Best Local Similarity 47.1%; Pred. No. 4.5;
Matches 106; Conservative 15; Mismatches 104; Indels 0; Gaps 0;
QY 23 AAAAAACCGCTAAGAAATCAGAGTTAAAAATCAGAAATATATCATTTCTTGA 82
DB 527 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 468
QY 83 CACATACATGTTCTTTTATACAAAAATATACACACAAATATTAATGTTGTAT 142
DB 467 AAAAAATATTAATATTTTATATATAATAAATWAAATGAATATATAATATAA 408
QY 143 TAGTGCTTATATAATATAATCTAGGGTTTACGCAACAATTTATCGATAATAA 202
DB 407 TTTTATAAAAAATAAANAATTAATTAATTAATAAATAAANAANAANAANAATA 348
QY 203 ATACTTTTACGCAACAATTTATAGACGGGAGGAATACCA 247
DB 347 ATAATWAAATAATATTWAAAAAANAANAANAANAANAANAANAANAANA 303
RESULT 14
BX458623/c 1201 bp mRNA linear EST 22-MAY-2003
LOCUS BX458623 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE006YG03
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX458623
VERSION BX458623.1 GI:31019361
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.S., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9800.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE006AD02QP1&cluster=9800.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE006AD02QP1.
FEATURES
source Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE006YG03"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
ORIGIN
Query Match 21.1%; Score 52.6; DB 13; Length 1201;
Best Local Similarity 41.9%; Pred. No. 3.7;
Matches 91; Conservative 30; Mismatches 96; Indels 0; Gaps 0;
QY 15 ATTTCGGAATAAACCGCTAAGAAATCAGAGTTAAAAATCAGAAATATATCATTA 74
DB 1056 ATTGTTWRRGRARRAAATKWTATAAAAAANAANAANAANAANAANAANA 997
QY 75 TTCTTCACATACATGTTCTTTTATACAAAAAANAANAATACACAAATTAATTCG 134
DB 996 TTTTITTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 937
QY 135 TGTTGATTAGGTGTTATATAATAATCTAGGGGTTTACGCAACAATTTATCG 194
DB 936 TATTTTWTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 877
QY 195 ATAAATTAATCTTTTACGCAACAATTTATAGA 231
DB 876 TTWTAAAAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAA 840
RESULT 15
AU074022
LOCUS AU074022 Dictyostelium discoideum ss (H.Urushihara) Dictyostelium
DEFINITION AU074022 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSJ301, mRNA sequence.
ACCESSION AU074022
VERSION AU074022.1 GI:5180443
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 240)
AUTHORS Urushihara, H.
TITLE Developmental cDNA in Dictyostelium discoideum (1999)
JOURNAL Unpublished (1999)
COMMENT Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
Tel: 81-298-53-4664
Fax: 81-298-53-6614
Email: hideko@biol.tsukuba.ac.jp
PROJECT = Dictyostelium discoideum cDNA project in Japan.
FEATURES
source Location/Qualifiers

```

source
1..240
/organism="Dictyostelium discoideum"
/mol_type="mRNA"
/strain="AX4"
/db_xref="taxon:44689"
/clone="SSJ301"
/dev_stage="slug"
/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"

ORIGIN

Query Match      21.0%; Score 52.4; DB 9; Length 240;
Best Local Similarity 50.9%; Pred. No. 9.1; Mismatches 109; Indels 0; Gaps 0;
Matches 113; Conservative 0;

QY 22 AAAAAACCCAGCGTAAAGAAATCAGAGTTAAAGAAATCAGAGAAATATATATATTCCTTG 81
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 7  AAAAAAAAAAAAAAAAAATTCATATAAAAAAAAAAAAAAAAAATTCACATATAATA 66
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 82 ACACATACATGTTCTTTTATACAAAAAATAACACACACATATTAATTGTTGTTGTA 141
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 67 AAAAAATTTTTCCTTTATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 126
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 142 TTAGGTGTTATATAATAATAATCTAGGGTTTAAACGCAACACAAATTTATCGATAAATA 201
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 127 TAAATAATATATATANNNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 186
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 202 AATACTTTGTAGCGCAACACAAATTTATAGACGCGGAGGAAAT 243
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 187 AATGGGTAAAAAGATAAAAAATGTATCAGTTGAAGAAGAGT 228
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Search completed: February 28, 2004, 09:05:55
Job time : 2941.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 05:48:29 ; Search time 171.373 Seconds
(without alignments)
2925.120 Million cell updates/sec

Title: US-09-895-435A-2
Perfect score: 118
Sequence: 1 taatacaacacataattat.....agacgaacacacatttatag 118

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	118	100.0	118	7 ABZ75336	Abz75336 B. thurin
2	118	100.0	244	7 ABZ75335	Abz75335 B. thurin
3	118	100.0	249	7 ABZ75338	Abz75338 B. thurin
4	114.4	96.9	116	7 ABZ75341	Abz75341 B. thurin
5	114.4	96.9	116	7 ABZ75340	Abz75340 B. thurin
6	64	54.2	116	7 ABZ75342	Abz75342 B. thurin
7	51	43.2	124	7 ABZ75343	Abz75343 B. thurin
8	41.8	35.4	2000	7 ADA73029	Ada73029 Rice gene
9	39.2	33.2	17594	6 ABL34026	Ab134026 Human imm
10	38	32.2	750	7 ACF70232	Acf70232 Photorhab
11	38	32.2	110000	7 ACF65385_3	Continuation (4 of
12	38	32.2	110000	7 ACF65385_4	Continuation (5 of
13	38	32.2	110000	7 ACF67367_32	Continuation (33 o
14	38	32.2	110000	7 ACF67367_33	Continuation (34 o
15	37.6	31.9	17594	6 ABL34027	Ab134027 Human imm
16	37.4	31.7	7008	4 AAS46531	Aas46531 Tumour su
17	37	31.4	14307	6 ABL32729	Ab132729 Human imm
18	36.4	30.8	4421	4 AAK76448	Aak76448 Human imm
19	36.4	30.8	4421	4 AAK76449	Aak76449 Human imm
20	35.4	30.0	230	3 AAA32056	Aaa32056 Plant mic
21	35.2	29.8	5349	6 ABL32340	Ab132340 Human imm
22	34.4	29.2	2778	6 AAD48244	Aad48244 Ehrlichia
23	34.4	29.2	6651	4 ABL03956	Ab103956 Drosophil

C 24	34.4	29.2	7889	4	ABL03922	Ab103922 Drosophil
C 25	34	28.8	5883	6	ABL33732	Ab133732 Human imm
C 26	34	28.8	17211	6	ABL32655	Ab132655 Human imm
C 27	34	28.8	32236	4	AAK91250	Aak91250 Human dig
C 28	34	28.8	110000	6	ABA03041_13	Continuation (14 o
C 29	33.8	28.6	2430	9	ADC49268	Adc49268 cDNA enco
C 30	33.8	28.6	5963	6	ABL34613	Ab134613 Human met
C 31	33.8	28.6	19787	6	ABL33450	Ab133450 Human imm
C 32	33.8	28.6	34548	6	ABL70603	Ab170603 Chemicall
C 33	33.6	28.5	244	7	ABZ75335	Abz75335 B. thurin
C 34	33.6	28.5	249	7	ABZ75338	Abz75338 B. thurin
C 35	33.6	28.5	6181	4	ABL21156	Ab121156 Drosophil
C 36	33.6	28.5	6388	4	ABL21154	Ab121154 Drosophil
C 37	33.6	28.5	6388	4	ABL13138	Ab113138 Drosophil
C 38	33.6	28.5	6392	4	ABL21158	Ab121158 Drosophil
C 39	33.6	28.5	14353	6	ABL34435	Ab134435 Human imm
C 40	33.6	28.5	15649	6	ABL70543	Ab170543 Chemicall
C 41	33.4	28.3	6113	6	ABL32431	Ab132431 Human imm
C 42	33.4	28.3	6113	6	ABL92205	Ab192205 Chemicall
C 43	33.4	28.3	6113	6	ABL49314	Ab149314 Human pol
C 44	33.2	28.1	306	3	AAC93751	Aac93751 Cat flea
C 45	33.2	28.1	700	4	AAH93025	Aah93025 Human inf

ALIGNMENTS

RESULT 1
ABZ75336
ID ABZ75336 standard; DNA; 118 BP.
XX
AC ABZ75336;
DT 06-MAY-2003 (first entry)
DE B. thuringiensis TnpI recognition target sequence TRT'.
KW TnpI; gene therapy; recombinase; TRT; TRT'; genetic engineering;
KW TnpI recognised target; TRT'; ds.
OS Bacillus thuringiensis.
FH Key Location/Qualifiers
FT repeat_region 1..32 /*tag= a
FT /*rpt_type= INVERTED
FT repeat_unit 1..14 /*tag= b
FT repeat_region 65..116 /*tag= c
FT repeat_unit 65..80 /*rpt_type= DIRECT
FT /*tag= d
WO2003004652-A2.
16-JAN-2003.
28-JUN-2002; 2002WO-EP007176.
30-JUN-2001; 2001US-00895435.
(EUMO-) EURO MOLECULAR BIOLOGY LAB.
(UYLO-) UNIV CATHOLIQUE LOUVAIN.
Stewart FA, Zhang Y, Hallett B;
WPI; 2003-221599/21.
A composition comprising recombinase TnpI and the recombination substrate TRT or TRT' sequences, useful for genetic engineering, TnpI-mediated genetic recombination, or for gene therapy.

Claim 6; Page 55; 58pp; English.

The invention relates to a novel composition comprising an isolated DNA molecule comprising TnpI recombination substrates. The composition may have a use in gene therapy, and as a vaccine. The recombinase TnpI and the recombination substrates TRT or TRT' sequences are useful for genetic engineering. The kits are useful for TnpI-mediated genetic recombination or site-specific recombination. The present sequence represents the TnpI recognised target (TRT) TRT'.

Sequence 118 BP; 49 A; 14 C; 15 G; 40 T; 0 U; 0 Other;

Query Match 100.0%; Score 118; DB 7; Length 118;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAATACAGCACAAATTATTCGATAAATAAATCGTTAGTGTTATTAATAAATCTAGGGG 60
DB 1 TAATACAGCACAAATTATTCGATAAATAAATCGTTAGTGTTATTAATAAATCTAGGGG 60

QY 61 TTATACGCACCAACAATTTATCGATAAATAAATCTTTTAGCGCAGCACAAATTTATAG 118
DB 61 TTATACGCACCAACAATTTATCGATAAATAAATCTTTTAGCGCAGCACAAATTTATAG 118

RESULT 2
ABZ75335
ID ABZ75335 standard; DNA; 244 BP.
AC ABZ75335;
XX DT 06-MAY-2003 (first entry)
DE B. thuringiensis TnpI recognition target sequence TRT'''.
DE TnpI; gene therapy; recombinase; TRT; TRT'; genetic engineering;
KW TnpI recognised target; TRT''; ds.
XX Bacillus thuringiensis.

Key Location/Qualifiers
repeat_region 37..59 /tag= a /rpt_type= DIRECT
repeat_unit 37..45 /tag= b /rpt_type= INVERTED
repeat_region 111..142 /tag= c /rpt_type= DIRECT
repeat_unit 111..124 /tag= d /rpt_type= INVERTED
repeat_region 175..211 /tag= e /rpt_type= DIRECT
repeat_unit 175..190 /tag= f /rpt_type= INVERTED

WO2003004652-A2.
16-JAN-2003.
28-JUN-2002; 2002WO-EP007176.
30-JUN-2001; 2001US-00895435.
(EUMO-) EURO MOLECULAR BIOLOGY LAB.
(UYLO-) UNIV CATHOLIQUE LOUVAIN.
Stewart FA, Zhang Y, Hallett B;


```
FT repeat_region 65..116
FT /tag= c
FT /rpt_type= DIRECT
FT repeat_unit 65..180
FT /tag= d
FT WO2003004652-A2.
XX PN
XX XX
XX PD 16-JAN-2003.
XX XX
XX PF 28-JUN-2002; 2002WO-EP007176.
XX PR 30-JUN-2001; 2001US-00895435.
XX PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX PI Stewart FA, Zhang Y, Hallet B;
XX XX
XX DR WPI; 2003-221599/21.
XX XX
XX PT A composition comprising recombinase Tnpi and the recombination substrate
XX TRT or TRT' sequences, useful for genetic engineering, Tnpi-mediated
XX PT genetic recombination, or for gene therapy.
XX PS Disclosure; Page 12; 58pp; English.
XX XX
XX CC The invention relates to a novel composition comprising an isolated DNA
XX CC molecule comprising Tnpi recombination substrates. The composition may
XX CC have a use in gene therapy, and as a vaccine. The recombinase Tnpi and
XX CC the recombination substrates TRT or TRT' sequences are useful for genetic
XX CC engineering. The kits are useful for Tnpi-mediated genetic recombination.
XX CC The Tnpi-mediated site-specific recombination is useful for deleting or
XX CC inserting DNA sequences to modulate gene expression, integrating new
XX CC nucleotide sequences into a DNA molecule, such as an episome or genomic
XX CC DNA, for gene therapy, or for producing food-grade genetically modified
XX CC organisms. The present sequence represents a Tnpi recognised target (TRT)
XX CC functional variant TRT'.
XX SQ Sequence 116 BP; 47 A; 14 C; 14 G; 41 T; 0 U; 0 Other;
Query Match 96.9%; Score 114.4; DB 7; Length 116;
Best Local Similarity 99.1%; Pred. No. 3.1e-18;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TAATACACACATATTAATTTGTTGTTAGTGTATATTAATAATAATTAATCTAGGGG 60
Db 1 TAATACACACATATTAATTTGTTGTTAGTGTATATTAATAATAATATATCTAGGGG 60
QY 61 TTTAAGCGACACAAATTTATCGATAAATAAATACTTTTAGACGCAACACAAATTTAT 116
Db 61 TTTAAGCGACACAAATTTATCGATAAATAAATACTTTTAGACGCAACACAAATTTAT 116
RESULT 6
ABZ75342
ID ABZ75342 standard; DNA; 116 BP.
XX AC
XX ABZ75342;
XX DT
XX DT 06-MAY-2003 (first entry)
XX XX
XX DE B. thuringiensis Tnpi recognition target TRT' functional variant #3.
XX KW Tnpi; gene therapy; recombinase; TRT; TRT'; genetic engineering;
XX KW Tnpi recognised target; TRT'; db.
XX OS Bacillus thuringiensis.
XX XX
XX FH Key Location/Qualifiers
XX FT repeat_region 1..32
XX FT /tag= a
XX FT /rpt_type= INVERTED
```

```
FT repeat_unit 1..14
FT /tag= b
FT repeat_region 65..116
FT /tag= c
FT /rpt_type= DIRECT
FT repeat_unit 65..180
FT /tag= d
FT WO2003004652-A2.
XX PN
XX XX
XX PD 16-JAN-2003.
XX XX
XX PF 28-JUN-2002; 2002WO-EP007176.
XX PR 30-JUN-2001; 2001US-00895435.
XX PA (EUMO-) EURO MOLECULAR BIOLOGY LAB.
XX PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX PI Stewart FA, Zhang Y, Hallet B;
XX XX
XX DR WPI; 2003-221599/21.
XX XX
XX PT A composition comprising recombinase Tnpi and the recombination substrate
XX TRT or TRT' sequences, useful for genetic engineering, Tnpi-mediated
XX PT genetic recombination, or for gene therapy.
XX PS Disclosure; Page 13; 58pp; English.
XX XX
XX CC The invention relates to a novel composition comprising an isolated DNA
XX CC molecule comprising Tnpi recombination substrates. The composition may
XX CC have a use in gene therapy, and as a vaccine. The recombinase Tnpi and
XX CC the recombination substrates TRT or TRT' sequences are useful for genetic
XX CC engineering. The kits are useful for Tnpi-mediated genetic recombination.
XX CC The Tnpi-mediated site-specific recombination is useful for deleting or
XX CC inserting DNA sequences to modulate gene expression, integrating new
XX CC nucleotide sequences into a DNA molecule, such as an episome or genomic
XX CC DNA, for gene therapy, or for producing food-grade genetically modified
XX CC organisms. The present sequence represents a Tnpi recognised target (TRT)
XX CC functional variant TRT'.
XX SQ Sequence 116 BP; 27 A; 11 C; 5 G; 21 T; 0 U; 52 Other;
Query Match 54.2%; Score 64; DB 7; Length 116;
Best Local Similarity 55.2%; Pred. No. 2.6e-06;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 TAATACACACATATTAATTTGTTGTTAGTGTATATTAATAATAATTAATCTAGGGG 60
Db 1 TAATACACACATATTAATTTGTTGTTAGTGTATATTAATAATAATAATTTAT 116
QY 61 TTTAAGCGACACAAATTTATCGATAAATAAATACTTTTAGACGCAACACAAATTTAT 116
Db 61 TTTAAGCGACACAAATTTATCGATAAATAAATACTTTTAGACGCAACACAAATTTAT 116
RESULT 7
ABZ75343
ID ABZ75343 standard; DNA; 124 BP.
XX AC
XX ABZ75343;
XX DT
XX DT 06-MAY-2003 (first entry)
XX XX
XX DE B. thuringiensis Tnpi recognition target TRT consensus.
XX KW Tnpi; gene therapy; recombinase; TRT; TRT'; genetic engineering;
XX KW Tnpi recognised target; TRT'; db; consensus.
XX OS Bacillus thuringiensis.
XX XX
XX FH Key Location/Qualifiers
XX FT repeat_region 1..32
```

FT /*tag= a
FT /rpt_type= INVERTED
FT 1..14
FT /*tag= b
FT repeat_region 75..124
FT /*tag= c
FT /rpt_type= DIRECT
FT 75..89
FT /*tag= d

WO2003004652-A2.

16-JAN-2003.

28-JUN-2002; 2002WO-EP007176.

30-JUN-2001; 2001US-00895435.

(EUMO-) EURO MOLECULAR BIOLOGY LAB.

(UYLO-) UNIV CATHOLIQUE LOUVAIN.

Stewart FA, Zhang Y, Hallet B;

WPI; 2003-221599/21.

PS A composition comprising recombinase Tnpi and the recombination substrate
FT Trr or Trr' sequences, useful for genetic engineering, Tnpi-mediated
PT genetic recombination, or for gene therapy.

PS Disclosure; Page 13; 58pp; English.

CC The invention relates to a novel composition comprising an isolated DNA
CC molecule comprising Tnpi recombination substrates. The composition may
CC have a use in gene therapy, and as a vaccine. The recombinase Tnpi and
CC the recombination substrates Trr or Trr' sequences are useful for genetic
CC engineering. The kits are useful for Tnpi-mediated genetic recombination.
CC The Tnpi-mediated site-specific recombination is useful for deleting or
CC inserting DNA sequences to modulate gene expression, integrating new
CC nucleotide sequences into a DNA molecule, such as an episome or genomic
CC DNA, for gene therapy, or for producing food-grade genetically modified
CC organisms. The present sequence represents a Tnpi recognised target (Trr)
CC consensus sequence

SQ Sequence 124 BP; 36 A; 13 C; 7 G; 26 T; 0 U; 42 Other;

Query Match 43.2%; Score 51; DB 7; Length 124;

Best Local Similarity 85.6%; Pred. No. 0.003;

Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps 2;

QY 1 TAATACACACACATATTAATTTGTTGTTATTA-----GGTGTATAATAATATA 50

DB 1 TAATACACACACATATTAATTTGTTGTTATTAANNNNNNNNNNNNNNNNNNNNNNNN 60

QY 51 AATCTAGGGTTTAACGCAACACAAATTTATCGATAATAAATCTTTTAGACGCAACACA 110

DB 61 NNNNNNNNNNNNNNACGCAACACAAATTTA-CGATAATAAATCTTTTAGACGCAACACA 119

QY 111 ATTTA 115

DB 120 ATTTA 124

RESULT 8

ADA73029/c

ID ADA73029 standard; DNA; 2000 BP.

XX AC ADA73029;

XX 20-NOV-2003 (first entry)

DT Rice gene, SEQ ID 6355.

DE Plant; bacterial infection; fungal infection; viral infection; rice;

XX

KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 27; SEQ ID NO 6355; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 2000 BP; 499 A; 454 C; 403 G; 636 T; 0 U; 8 Other;

Query Match 35.4%; Score 41.8; DB 7; Length 2000;

Best Local Similarity 59.8%; Pred. No. 0.48;

Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 TAATACACACACATATTAATTTGTTGTTATTAATAATAATAATAATACTAGGGG 60

DB 633 TAATACACACACACATATTTATTAATAATAGTTTAAAGCTAATTTCAATTT 574

QY 61 TTTAAGCAACACAAATTTATCGATAATAATACTTTTAGACGCAACACAAATTTATA 117

DB 573 TTAGATGCTGTTAAATTTTCTATATTTAAATAGGAAATAATCAAAATAACTTATA 517

RESULT 9

ABL34026

ID ABL34026 standard; DNA; 17594 BP.

XX ABL34026;

XX 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 1999.

XX Human; immune system disease; cytosine methylation; antiasthmatic;

XX antiarteriosclerotic; antianaemic; cytostatic; neotropic;

XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;

XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;

XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;

XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;

XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

XX Homo sapiens.

XX

RESULT 13
ACF67367 32/c
Continuation (33 of 57) of ACF67367 from base 3200001 (Photorhabdus luminescens nucleoti
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000

WF	ACF67367_07	700001	810000
WF	ACF67367_08	800001	910000
WF	ACF67367_09	900001	1010000
WF	ACF67367_10	1000001	1110000
WF	ACF67367_11	1100001	1210000

WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000

WE	ACF67367_24	2400001	2410000
WE	ACF67367_25	2500001	2610000
WE	ACF67367_26	2600001	2710000
WE	ACF67367_27	2700001	2810000
WE	ACF67367_28	2800001	2910000

WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400011	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5710000
WP	ACF67367_57	5700001	5810000
WP	ACF67367_58	5800001	5910000
WP	ACF67367_59	5900001	6010000
WP	ACF67367_60	6000001	6110000
WP	ACF67367_61	6100001	6210000
WP	ACF67367_62	6200001	6310000
WP	ACF67367_63	6300001	6410000
WP	ACF67367_64	6400001	6510000
WP	ACF67367_65	6500001	6610000
WP	ACF67367_66	6600001	6710000
WP	ACF67367_67	6700001	6810000
WP	ACF67367_68	6800001	6910000
WP	ACF67367_69	6900001	7010000
WP	ACF67367_70	7000001	7110000
WP	ACF67367_71	7100001	7210000
WP	ACF67367_72	7200001	7310000
WP	ACF67367_73	7300001	7410000
WP	ACF67367_74	7400001	7510000
WP	ACF67367_75	7500001	7610000
WP	ACF67367_76	7600001	7710000
WP	ACF67367_77	7700001	7810000
WP	ACF67367_78	7800001	7910000
WP	ACF67367_79	7900001	8010000
WP	ACF67367_80	8000001	8110000
WP	ACF67367_81	8100001	8210000
WP	ACF67367_82	8200001	8310000
WP	ACF67367_83	8300001	8410000
WP	ACF67367_84	8400001	8510000
WP	ACF67367_85	8500001	8610000
WP	ACF67367_86	8600001	8710000
WP	ACF67367_87	8700001	8810000
WP	ACF67367_88	8800001	8910000
WP	ACF67367_89	8900001	9010000
WP	ACF67367_90	9000001	9110000
WP	ACF67367_91	9100001	9210000
WP	ACF67367_92	9200001	9310000
WP	ACF67367_93	9300001	9410000
WP	ACF67367_94	9400001	9510000
WP	ACF67367_95	9500001	9610000
WP	ACF67367_96	9600001	9710000
WP	ACF67367_97	9700001	9810000
WP	ACF67367_98	9800001	9910000
WP	ACF67367_99	9900001	1000000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 108766 GCAAATCAAGTAAATAGATAAAGAAATTCCTTTAGAAACAGCA 108725

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000

WP	ACF67367_07	700001	810000
WP	ACF67367_08	900001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000

WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000

ACF67367_24 2310000
 ACF67367_25 2500001
 ACF67367_26 2610000
 ACF67367_27 2710000
 ACF67367_28 2800001
 ACF67367_29 2910000

WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5710000
WP	ACF67367_57	5700001	5810000
WP	ACF67367_58	5800001	5910000
WP	ACF67367_59	5900001	6010000
WP	ACF67367_60	6000001	6110000
WP	ACF67367_61	6100001	6210000
WP	ACF67367_62	6200001	6310000
WP	ACF67367_63	6300001	6410000
WP	ACF67367_64	6400001	6510000
WP	ACF67367_65	6500001	6610000
WP	ACF67367_66	6600001	6710000
WP	ACF67367_67	6700001	6810000
WP	ACF67367_68	6800001	6910000
WP	ACF67367_69	6900001	7010000
WP	ACF67367_70	7000001	7110000
WP	ACF67367_71	7100001	7210000
WP	ACF67367_72	7200001	7310000
WP	ACF67367_73	7300001	7410000
WP	ACF67367_74	7400001	7510000
WP	ACF67367_75	7500001	7610000
WP	ACF67367_76	7600001	7710000
WP	ACF67367_77	7700001	7810000
WP	ACF67367_78	7800001	7910000
WP	ACF67367_79	7900001	8010000
WP	ACF67367_80	8000001	8110000
WP	ACF67367_81	8100001	8210000
WP	ACF67367_82	8200001	8310000
WP	ACF67367_83	8300001	8410000
WP	ACF67367_84	8400001	8510000
WP	ACF67367_85	8500001	8610000
WP	ACF67367_86	8600001	8710000
WP	ACF67367_87	8700001	8810000
WP	ACF67367_88	8800001	8910000
WP	ACF67367_89	8900001	9010000
WP	ACF67367_90	9000001	9110000
WP	ACF67367_91	9100001	9210000
WP	ACF67367_92	9200001	9310000
WP	ACF67367_93	9300001	9410000
WP	ACF67367_94	9400001	9510000
WP	ACF67367_95	9500001	9610000
WP	ACF67367_96	9600001	9710000
WP	ACF67367_97	9700001	9810000
WP	ACF67367_98	9800001	9910000
WP	ACF67367_99	9900001	1000000

WP	ACF67367_37	3700001	3510000
WP	ACF67367_38	3800001	3910000

```
WP ACP67367_39 3900001 4010000
WP ACP67367_40 4000001 4110000
WP ACP67367_41 4100001 4210000
WP ACP67367_42 4200001 4310000
WP ACP67367_43 4300001 4410000
WP ACP67367_44 4400001 4510000
WP ACP67367_45 4500001 4610000
WP ACP67367_46 4600001 4710000
WP ACP67367_47 4700001 4810000
WP ACP67367_48 4800001 4910000
WP ACP67367_49 4900001 5010000
WP ACP67367_50 5000001 5110000
WP ACP67367_51 5100001 5210000
WP ACP67367_52 5200001 5310000
WP ACP67367_53 5300001 5410000
WP ACP67367_54 5400001 5510000
WP ACP67367_55 5500001 5610000
WP ACP67367_56 5600001 5648894

Query Match 32.2%; Score 38; DB 7; Length 110000;
Best Local Similarity 60.8%; Pred. No. 4;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 7 AACACATATTATTCGTTGTTAGTGTATATATAAATAAATCTAGGGGTTTAAAC 66
Db 8826 AAAAAATTGATTTGAAATCGGAAGAGATTCATCAGAAATGTTATGATACAAATATT 8767
Qy 67 GCAACACAAATTTATCGATAAATAAATACTTTTAGACGCAACA 108
Db 8766 GCAATCAAGTATAGATAAGAAATCTTTTAGAACAGCA 8725

Query Match 31.9%; Score 37.6; DB 6; Length 17594;
Best Local Similarity 57.8%; Pred. No. 4.8;
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATATTAAATGTTGTTAGTGTATATAAATAAATAAATACTAGGGG 60
Db 16808 TATATATTAACTATAAATACTATTTAACTATATAAATAAATAAATAAATAAATA 16749
Qy 61 TTTAACGCAACACAAATTTATCGATAAATAAATACTTTTAGACGCAACAAATTTAT 116
Db 16748 TATATATCAATTTAAATATATATAAATAAATAAATAAATAAATAAATAAATA 16693

Search completed: February 28, 2004, 07:22:24
Job time : 173.373 secs
```

```
RESULT 15
ABL34027/c
ID ABL34027 standard; DNA; 17594 BP.
XX
AC ABL34027;
XX
DT 26-MAR-2002 (first entry)
DE
DE Human immune system associated gene SEQ ID NO: 2000.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
XX WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 2000; 32pp + Sequence Listing; German.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:07:41 ; Search time 130.218 Seconds
(without alignments)
3268.988 Million cell updates/sec

Title: US-09-895-435A-2

Perfect score: 118

Sequence: 1 taatacaacaataataat.....agacgaacacaatttatag 118

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2353733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	118	10	US-09-895-435-2
2	118	100.0	244	10	US-09-895-435-1
3	118	100.0	249	10	US-09-895-435-4
4	114.4	96.9	116	10	US-09-895-435-7
5	114.4	96.9	116	10	US-09-895-435-8
6	64	54.2	116	10	US-09-895-435-9
7	51	43.2	124	10	US-09-895-435-10
8	41.6	35.3	761	12	US-10-424-599-137579
9	35.2	33.2	17594	14	US-10-311-455-1999
10	37.6	31.9	17594	14	US-10-311-455-2000
11	37	31.4	14307	14	US-10-311-455-702
12	35.2	29.8	5349	14	US-10-311-455-313
13	35	29.7	497	12	US-10-424-599-110450
14	34.4	29.2	2778	14	US-10-081-051-23
15	34.4	29.2	96898	15	US-10-417-476-3

C 16	34	28.8	5883	14	US-10-311-455-1705	Sequence 1705, Ap
C 17	34	28.8	17211	14	US-10-311-455-628	Sequence 628, App
C 18	33.8	28.6	5963	14	US-10-340-485-166	Sequence 166, App
C 19	33.8	28.6	19787	14	US-10-311-455-1423	Sequence 1423, Ap
C 20	33.6	28.5	244	10	US-09-895-435-1	Sequence 1, Appli
C 21	33.6	28.5	249	10	US-09-895-435-4	Sequence 4, Appli
C 22	33.6	28.5	600	15	US-10-027-632-222461	Sequence 222461,
C 23	33.6	28.5	14353	14	US-10-311-455-2408	Sequence 2408, Ap
C 24	33.4	28.3	6113	14	US-10-204-708-14	Sequence 14, Appl
C 25	33.4	28.3	6113	14	US-10-311-455-404	Sequence 404, App
C 26	33.2	28.1	306	10	US-09-991-936-246	Sequence 246, App
C 27	33.2	28.1	6185	14	US-10-311-455-386	Sequence 386, App
C 28	33	28.0	16236	14	US-10-311-455-995	Sequence 995, App
C 29	32.8	27.8	433	14	US-10-301-997-64	Sequence 64, Appl
C 30	32.8	27.8	5683	14	US-10-311-455-1780	Sequence 1780, Ap
C 31	32.8	27.8	5683	14	US-10-240-485-144	Sequence 144, App
C 32	32.8	27.8	6072	14	US-10-311-455-3	Sequence 3, Appli
C 33	32.8	27.8	17211	14	US-10-311-455-627	Sequence 627, App
C 34	32.6	27.6	2000	9	US-09-938-842A-5054	Sequence 5054, Ap
C 35	32.6	27.6	2000	11	US-09-938-842A-5054	Sequence 5054, Ap
C 36	32.6	27.6	6112	14	US-10-311-455-462	Sequence 462, App
C 37	32.6	27.6	19082	14	US-10-311-455-599	Sequence 599, App
C 38	32.4	27.5	2019	9	US-09-792-630-36	Sequence 36, Appl
C 39	32.4	27.5	2019	10	US-09-953-351-36	Sequence 36, Appl
C 40	32.4	27.5	2019	13	US-10-080-376-36	Sequence 36, Appl
C 41	32.4	27.5	2019	14	US-10-082-671-42	Sequence 42, Appl
C 42	32.4	27.5	2019	14	US-10-097-100-36	Sequence 36, Appl
C 43	32.4	27.5	2019	14	US-10-023-208-36	Sequence 36, Appl
C 44	32.4	27.5	7498	14	US-10-311-455-230	Sequence 230, App
C 45	32.4	27.5	108316	15	US-10-292-798-1789	Sequence 1789, Ap

ALIGNMENTS

RESULT 1
US-09-895-435-2
; Sequence 2, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallett, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; TYPE: DNA
; LENGTH: 118
; ORGANISM: Bacillus thuringiensis
US-09-895-435-2

Query Match 100.0%; Score 118; DB 10; Length 118;
Best Local Similarity 100.0%; Pred. No. 3.1e-18;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TAATAACAACAAATTAATTCGTGTGTTAGCTTATTAATAATAATAATCTAGGGG 60
Db 1 TAATAACAACAAATTAATTCGTGTGTTAGCTTATTAATAATAATAATCTAGGGG 60
Qy 61 TTTTACGCAACAACTTTATCGATAATAATAATCTTTTAGACGCAACAATTATAG 118
Db 61 TTTTACGCAACAACTTTATCGATAATAATAATCTTTTAGACGCAACAATTATAG 118

RESULT 2
US-09-895-435-1
; Sequence 1, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:

```

; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-7

Query Match          96.9%; Score 114.4; DB 10; Length 116;
Best Local Similarity 99.1%; Pred. No. 2.1e-17;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TAATACGACACATATTAATTCGTTGTTATTAGGCTGTATATAAATAAATCTAGGGG 60
DB      1 TAATACGACACATATTAATTCGTTGTTATTAGGCTGTATATAAATAAATCTAGGGG 60

QY      61 TTTAACGCAACACAATTTATCGATAAATAAATACITTTTAGCGCAACACAATTTAT 116
DB      61 TTTAACGCAACACAATTTATCGATAAATAAATACITTTTAGCGCAACACAATTTAT 116

RESULT 5
US-09-895-435-8
; Sequence 8, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-8

Query Match          96.9%; Score 114.4; DB 10; Length 116;
Best Local Similarity 99.1%; Pred. No. 2.1e-17;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TAATACGACACATTAATTCGTTGTTATTAGGCTGTATATAAATAAATCTAGGGG 60
DB      1 TAATACGACACATTAATTCGTTGTTATTAGGCTGTATATAAATAAATCTAGGGG 60

QY      61 TTTAACGCAACACAATTTATCGATAAATAAATACITTTTAGCGCAACACAATTTAT 116
DB      61 TTTAACGCAACACAATTTATCGATAAATAAATACITTTTAGCGCAACACAATTTAT 116

RESULT 6
US-09-895-435-9
; Sequence 9, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

```

```
; NAME/KEY: N region
; LOCATION: 33 .. 64
; OTHER INFORMATION: n = a, t, g, or c
; NAME/KEY: N_region
; LOCATION: 81 .. 100
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-9

Query Match          54.2%; Score 64; DB 10; Length 116;
Best Local Similarity 55.2%; Pred. No. 8.2e-06;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 TAATACACACAATATTAATTTGTTGTTATAGTGTATATAATAATAATACTAGGGG 60
Db 1 TAATACACACAATATTAATTTGTTGTTATANNNNNNNNNNNNNNNNNNNNNNNNNN 60
QY 61 TTTAACGCGACACAATTTATCGATAATAATAACTTTTAGCGCGACACAATTTAT 116
Db 61 NNNNACGCAACAATTTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNACGCAACAATTTAT 116

RESULT 7
US-09-895-435-10
; Sequence 10, Application US/09895435
; Publication No. US20030113837A1
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 10
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33 .. 74
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-10

Query Match          43.2%; Score 51; DB 10; Length 124;
Best Local Similarity 65.6%; Pred. No. 0.0082;
Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps 2;

QY 1 TAATACACACAATATTAATTTGTTGTTATTA-----GGTGTATAATAATATA 50
Db 1 TAATACACACAATATTAATTTGTTGTTATANNNNNNNNNNNNNNNNNNNNNNNNNN 60
QY 51 AATCTAGGGTTTAACGCAACAATTTATCGATAATAATAACTTTTAGCGCGACACA 110
Db 61 NNNNNNNNNNNNACGCAACAATTTA-CGATAATAATAACTTTTAGCGCGACACA 119
QY 111 ATTTA 115
Db 120 ATTTA 124

RESULT 8
US-10-424-599-137579
; Sequence 137579, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 137579
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95242C.1
US-10-424-599-137579

Query Match          35.3%; Score 41.6; DB 12; Length 761;
Best Local Similarity 62.5%; Pred. No. 1.9;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 8 ACACAATATTAATTTGTTGTTATAGTGTATATAATAATAATACTAGGGGTTTAA 67
Db 632 AAACCAATTTATATTTTGAATTAGAAGTGGGAATAACAAGATTACCGGTTTCCC 691
QY 68 CAACACAATTTATCGATAATAATAACTTTTAGACGCAACACAA 111
Db 692 CTTTTCATTTAATAATAAAAAAATTCATTTATGGCAATATA 735

RESULT 9
US-10-311-455-1999
; Sequence 1999, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Det
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1999
; LENGTH: 17594
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1999

Query Match          33.2%; Score 39.2; DB 14; Length 17594;
Best Local Similarity 58.6%; Pred. No. 15;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1 TAATACACACAATATTAATTTGTTGTTATAGTGTATATAATAATAATACTAGGG 60
Db 787 TAGTATATTAATTTATGGAATTTAAATTAATAATAATAATAATAATAATTTG 846
QY 61 TTTAACGCAACAATTTATCGATAATAATAACTTTTAGACGCAACACAATTTAT 116
Db 847 TATATATTAGTTTAAATATATATAATAATAATAATAATAATAAGTAAAGTTGT 902

RESULT 10
US-10-311-455-2000/c
; Sequence 2000, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
```



```
, LOCATION: (25983)..(25997)
, OTHER INFORMATION: BIVM Exon 2 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (25993)..(26119)
, OTHER INFORMATION: BIVM Exon 2
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (26110)..(26124)
, OTHER INFORMATION: BIVM Exon 2 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (30592)..(30606)
, OTHER INFORMATION: BIVM Exon 3 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (30602)..(30697)
, OTHER INFORMATION: BIVM Exon 3
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (30688)..(30702)
, OTHER INFORMATION: BIVM Exon 3 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31298)..(31312)
, OTHER INFORMATION: BIVM Exon 4 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31308)..(31412)
, OTHER INFORMATION: BIVM Exon 4
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31403)..(31417)
, OTHER INFORMATION: BIVM Exon 4 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31620)..(31634)
, OTHER INFORMATION: BIVM Exon 5 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31630)..(31724)
, OTHER INFORMATION: BIVM Exon 5
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (31715)..(31729)
, OTHER INFORMATION: BIVM Exon 5 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (41120)..(41134)
, OTHER INFORMATION: BIVM Exon 6 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (41130)..(41262)
, OTHER INFORMATION: BIVM Exon 6
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (41253)..(41267)
, OTHER INFORMATION: BIVM Exon 6 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (44021)..(44035)
, OTHER INFORMATION: BIVM Exon 7 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (44031)..(44117)
, OTHER INFORMATION: BIVM Exon 7
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (44108)..(44122)
, OTHER INFORMATION: BIVM Exon 7 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (48198)..(48212)
```

```
, OTHER INFORMATION: BIVM Exon 8 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (48208)..(48304)
, OTHER INFORMATION: BIVM Exon 8
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (48295)..(48309)
, OTHER INFORMATION: BIVM Exon 8 splice donor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (49127)..(49141)
, OTHER INFORMATION: BIVM Exon 9 splice acceptor site
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (49137)..(51096)
, OTHER INFORMATION: BIVM Exon 9
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (49428)..(49430)
, OTHER INFORMATION: BIVM translation termination codon (TGA)
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (50039)..(50248)
, OTHER INFORMATION: Alu sequence
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: (55216)..(55975)
, OTHER INFORMATION: ERCCs Exon 1
, FEATURE:
, NAME/KEY: misc_feature

Query Match 29.2%; Score 34.4; DB 15; Length 96898;
Best Local Similarity 59.0%; Pred. No. 3e-02; Indels 0; Gaps 0;
Matches 59; Conservative 0; Mismatches 41;

QY 16 TTAATTGTTGTTATTAGGTGTTTATAATAATAATAATACTAGCGGTTTACGCAACAA 75
DB 81258 TAAATTTGTTTAAAAATCTGTAAAAACAGATACATGTCGTCTTTTAAAAATTAAACAT 81199

QY 76 TTTATCGATAATAATACTTTTAGACGCAACACAATTTA 115
DB 81198 TTTATATATATAATAATACATATGTATGTAAAAACAATAA 81159
```

Search completed: February 28, 2004, 09:12:49
JOB time : 131.718 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:13:57 ; Search time 1856.17 Seconds
(without alignments)
2236.845 Million cell updates/sec

Title: US-09-895-435A-2

Perfect score: 118

Sequence: 1 taatacaacacaattattat.....agacgcaacacaattattatg 118

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main:
1: /cgn2_6/prodata/2/pna/pctus COMB.seq.*
2: /cgn2_6/prodata/2/pna/US06 COMB.seq.*
3: /cgn2_6/prodata/2/pna/US07 COMB.seq.*
4: /cgn2_6/prodata/2/pna/US08 COMB.seq.*
5: /cgn2_6/prodata/2/pna/US081 COMB.seq.*
6: /cgn2_6/prodata/2/pna/US082 COMB.seq.*
7: /cgn2_6/prodata/2/pna/US083 COMB.seq.*
8: /cgn2_6/prodata/2/pna/US084 COMB.seq.*
9: /cgn2_6/prodata/2/pna/US085 COMB.seq.*
10: /cgn2_6/prodata/2/pna/US086 COMB.seq.*
11: /cgn2_6/prodata/2/pna/US087 COMB.seq.*
12: /cgn2_6/prodata/2/pna/US088 COMB.seq.*
13: /cgn2_6/prodata/2/pna/US089 COMB.seq.*
14: /cgn2_6/prodata/2/pna/US090 COMB.seq.*
15: /cgn2_6/prodata/2/pna/US091 COMB.seq.*
16: /cgn2_6/prodata/2/pna/US092A COMB.seq.*
17: /cgn2_6/prodata/2/pna/US092B COMB.seq.*
18: /cgn2_6/prodata/2/pna/US093A COMB.seq.*
19: /cgn2_6/prodata/2/pna/US093B COMB.seq.*
20: /cgn2_6/prodata/2/pna/US094 COMB.seq.*
21: /cgn2_6/prodata/2/pna/US095A COMB.seq.*
22: /cgn2_6/prodata/2/pna/US095B COMB.seq.*
23: /cgn2_6/prodata/2/pna/US095C COMB.seq.*
24: /cgn2_6/prodata/2/pna/US095D COMB.seq.*
25: /cgn2_6/prodata/2/pna/US096A COMB.seq.*
26: /cgn2_6/prodata/2/pna/US096B COMB.seq.*
27: /cgn2_6/prodata/2/pna/US096C COMB.seq.*
28: /cgn2_6/prodata/2/pna/US096D COMB.seq.*
29: /cgn2_6/prodata/2/pna/US096E COMB.seq.*
30: /cgn2_6/prodata/2/pna/US097A COMB.seq.*
31: /cgn2_6/prodata/2/pna/US097B COMB.seq.*
32: /cgn2_6/prodata/2/pna/US097C COMB.seq.*
33: /cgn2_6/prodata/2/pna/US098A COMB.seq.*
34: /cgn2_6/prodata/2/pna/US098B COMB.seq.*
35: /cgn2_6/prodata/2/pna/US098C COMB.seq.*
36: /cgn2_6/prodata/2/pna/US098D COMB.seq.*
37: /cgn2_6/prodata/2/pna/US099A COMB.seq.*
38: /cgn2_6/prodata/2/pna/US099B COMB.seq.*
39: /cgn2_6/prodata/2/pna/US099C COMB.seq.*
40: /cgn2_6/prodata/2/pna/US099D COMB.seq.*
41: /cgn2_6/prodata/2/pna/US099E COMB.seq.*
42: /cgn2_6/prodata/2/pna/US099F COMB.seq.*
43: /cgn2_6/prodata/2/pna/US100A COMB.seq.*

44: /cgn2_6/prodata/2/pna/US100B COMB.seq.*
45: /cgn2_6/prodata/2/pna/US101A COMB.seq.*
46: /cgn2_6/prodata/2/pna/US101B COMB.seq.*
47: /cgn2_6/prodata/2/pna/US102A COMB.seq.*
48: /cgn2_6/prodata/2/pna/US102B COMB.seq.*
49: /cgn2_6/prodata/2/pna/US103A COMB.seq.*
50: /cgn2_6/prodata/2/pna/US103B COMB.seq.*
51: /cgn2_6/prodata/2/pna/US104A COMB.seq.*
52: /cgn2_6/prodata/2/pna/US104B COMB.seq.*
53: /cgn2_6/prodata/2/pna/US106 COMB.seq.*
54: /cgn2_6/prodata/2/pna/US107A COMB.seq.*
55: /cgn2_6/prodata/2/pna/US107B COMB.seq.*
56: /cgn2_6/prodata/2/pna/US6000 COMB.seq.*
57: /cgn2_6/prodata/2/pna/US6001 COMB.seq.*
58: /cgn2_6/prodata/2/pna/US6002 COMB.seq.*
59: /cgn2_6/prodata/2/pna/US6003 COMB.seq.*
60: /cgn2_6/prodata/2/pna/US6004 COMB.seq.*
61: /cgn2_6/prodata/2/pna/US6005 COMB.seq.*
62: /cgn2_6/prodata/2/pna/US6006 COMB.seq.*
63: /cgn2_6/prodata/2/pna/US6007 COMB.seq.*
64: /cgn2_6/prodata/2/pna/US6008 COMB.seq.*
65: /cgn2_6/prodata/2/pna/US6009 COMB.seq.*
66: /cgn2_6/prodata/2/pna/US6010 COMB.seq.*
67: /cgn2_6/prodata/2/pna/US6011 COMB.seq.*
68: /cgn2_6/prodata/2/pna/US6012 COMB.seq.*
69: /cgn2_6/prodata/2/pna/US6013 COMB.seq.*
70: /cgn2_6/prodata/2/pna/US6014 COMB.seq.*
71: /cgn2_6/prodata/2/pna/US6015 COMB.seq.*
72: /cgn2_6/prodata/2/pna/US6016 COMB.seq.*
73: /cgn2_6/prodata/2/pna/US6017 COMB.seq.*
74: /cgn2_6/prodata/2/pna/US6018 COMB.seq.*
75: /cgn2_6/prodata/2/pna/US6019 COMB.seq.*
76: /cgn2_6/prodata/2/pna/US6020 COMB.seq.*
77: /cgn2_6/prodata/2/pna/US6021 COMB.seq.*
78: /cgn2_6/prodata/2/pna/US6022 COMB.seq.*
79: /cgn2_6/prodata/2/pna/US6023A COMB.seq.*
80: /cgn2_6/prodata/2/pna/US6023B COMB.seq.*
81: /cgn2_6/prodata/2/pna/US6024 COMB.seq.*
82: /cgn2_6/prodata/2/pna/US6025 COMB.seq.*
83: /cgn2_6/prodata/2/pna/US6026 COMB.seq.*
84: /cgn2_6/prodata/2/pna/US6027 COMB.seq.*
85: /cgn2_6/prodata/2/pna/US6028 COMB.seq.*
86: /cgn2_6/prodata/2/pna/US6029 COMB.seq.*
87: /cgn2_6/prodata/2/pna/US6030 COMB.seq.*
88: /cgn2_6/prodata/2/pna/US6031 COMB.seq.*
89: /cgn2_6/prodata/2/pna/US6032 COMB.seq.*
90: /cgn2_6/prodata/2/pna/US6033 COMB.seq.*
91: /cgn2_6/prodata/2/pna/US6034 COMB.seq.*
92: /cgn2_6/prodata/2/pna/US6035 COMB.seq.*
93: /cgn2_6/prodata/2/pna/US6036 COMB.seq.*
94: /cgn2_6/prodata/2/pna/US6037 COMB.seq.*
95: /cgn2_6/prodata/2/pna/US6038 COMB.seq.*
96: /cgn2_6/prodata/2/pna/US6039 COMB.seq.*
97: /cgn2_6/prodata/2/pna/US6040 COMB.seq.*
98: /cgn2_6/prodata/2/pna/US6041 COMB.seq.*
99: /cgn2_6/prodata/2/pna/US6042 COMB.seq.*
100: /cgn2_6/prodata/2/pna/US6043 COMB.seq.*
101: /cgn2_6/prodata/2/pna/US6044 COMB.seq.*
102: /cgn2_6/prodata/2/pna/US6045 COMB.seq.*
103: /cgn2_6/prodata/2/pna/US6046 COMB.seq.*
104: /cgn2_6/prodata/2/pna/US6047 COMB.seq.*
105: /cgn2_6/prodata/2/pna/US6048 COMB.seq.*
106: /cgn2_6/prodata/2/pna/US6049 COMB.seq.*
107: /cgn2_6/prodata/2/pna/US6050 COMB.seq.*
108: /cgn2_6/prodata/2/pna/US6051 COMB.seq.*
109: /cgn2_6/prodata/2/pna/US6052 COMB.seq.*
110: /cgn2_6/prodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES


```
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435-8

Query Match          96.9%; Score 114.4; DB 36; Length 116;
Best Local Similarity 99.1%; Pred. No. 1.8e-16;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60
Db 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60

Qy 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116
Db 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116

RESULT 9
US-09-895-435A-7
; Sequence 7, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435A-7

Query Match          96.9%; Score 114.4; DB 36; Length 116;
Best Local Similarity 99.1%; Pred. No. 1.8e-16;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60
Db 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60

Qy 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116
Db 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116

RESULT 10
US-09-895-435A-8
; Sequence 8, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-895-435A-8

Query Match          96.9%; Score 114.4; DB 36; Length 116;
Best Local Similarity 99.1%; Pred. No. 1.8e-16;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60
Db 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60

Qy 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116
Db 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116

RESULT 11
US-09-895-435-9
; Sequence 9, Application US/09895435
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33..64
; OTHER INFORMATION: n = a, t, g, or c
; NAME/KEY: N region
; LOCATION: 81..100
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-9

Query Match          54.2%; Score 64; DB 36; Length 116;
Best Local Similarity 55.2%; Pred. No. 5.8e-05;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60
Db 1 TAATACACACAAATTAATTTGTTGTTATAGGTGTTATATAAATAAATAAATCTAGGGG 60

Qy 61 TTTAAGCGCAACAATTTATCGATAAATAAATAAATAAATAAATAAATAAATAAATTTAT 116
Db 61 NNNNACGCAACAATTTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 116

RESULT 12
US-09-895-435A-9
; Sequence 9, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 116
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33..64
; OTHER INFORMATION: n = a, t, g, or c
; NAME/KEY: N region
; LOCATION: 81..100
; OTHER INFORMATION: n = a, t, g, or c
```

US-09-895-435A-9

Query Match 54.2%; Score 64; DB 36; Length 116;
Best Local Similarity 55.2%; Pred. No. 5.8e-05;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 TAATACACACAAATTAATGTTGTTATAGCTTATATAATAATCTAGGGG 60
DB 1 TAATACACACAAATTAATGTTGTTATANNNNNNNNNNNNNNNNNNNNNN 60
QY 61 TTTTACGCGACACAAATTTATCGATAAATAATCTTTTAGACGCGACACAAATTTAT 116
DB 61 NNNNACGCGACACAAATTTANNNNNNNNNNNNNNNNNNNNNNNACGCGACAAATTTAT 116

RESULT 13

US-09-895-435-10
; Sequence 10, Application US/09895435
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: N region
; LOCATION: 33..74
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435-10

Query Match 43.2%; Score 51; DB 36; Length 124;
Best Local Similarity 65.6%; Pred. No. 0.055;
Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps 2;
QY 1 TAATACACACAAATTAATGTTGTTATTA-----GGTGTATAATAATATA 50
DB 1 TAATACACACAAATTAATGTTGTTATANNNNNNNNNNNNNNNNNNNNNN 60
QY 51 AATCTAGGGGTTTAAACGCGACACAAATTTATCGATAAATAATCTTTTAGACGCGACACAA 110
DB 61 NNNNNNNNNNNNACGCGACACAAATTTA-CGATAAATAATCTTTTAGACGCGACACAA 119
QY 111 ATTTA 115
DB 120 ATTTA 124

RESULT 14

US-09-895-435A-10
; Sequence 10, Application US/09895435A
; GENERAL INFORMATION:
; APPLICANT: Stewart, A. F.
; APPLICANT: Zhang, Y.
; APPLICANT: Hallet, B.
; TITLE OF INVENTION: A New Tyrosine Recombinase for Genetic Engineering
; FILE REFERENCE: 9882-012-999
; CURRENT APPLICATION NUMBER: US/09/895,435A
; CURRENT FILING DATE: 2001-06-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:

; NAME/KEY: N region
; LOCATION: 33..74
; OTHER INFORMATION: n = a, t, g, or c
US-09-895-435A-10

Query Match 43.2%; Score 51; DB 36; Length 124;
Best Local Similarity 65.6%; Pred. No. 0.055;
Matches 82; Conservative 0; Mismatches 32; Indels 11; Gaps 2;
QY 1 TAATACACACAAATTAATGTTGTTATTA-----GGTGTATAATAATATA 50
DB 1 TAATACACACAAATTAATGTTGTTATANNNNNNNNNNNNNNNNNNNNNN 60
QY 51 AATCTAGGGGTTTAAACGCGACACAAATTTATCGATAAATAATCTTTTAGACGCGACACAA 110
DB 61 NNNNNNNNNNACGCGACACAAATTTA-CGATAAATAATCTTTTAGACGCGACACAA 119
QY 111 ATTTA 115
DB 120 ATTTA 124

RESULT 15

US-09-887-272A-6355/c
; Sequence 6355, Application US/09887272A
; GENERAL INFORMATION:
; APPLICANT: Hou, Yu-Ming
; APPLICANT: Quan, Sheng
; APPLICANT: Chang, Hur-Song
; APPLICANT: Zhu, Tong
; APPLICANT: Whitham, Steve
; APPLICANT: Goff, Steve
; APPLICANT: Glazebrook, Jane
; APPLICANT: Chen, Wenguiang
; APPLICANT: Katagiri, Fumiaki
; APPLICANT: Xie, Zhiyi
; APPLICANT: Tao, Yi
; APPLICANT: Zou, Guangzhou
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: PLANT GENES INVOLVED IN DEFENSE AGAINST
; FILE REFERENCE: 1360.003US2
; CURRENT APPLICATION NUMBER: US/09/887,272A
; CURRENT FILING DATE: 2001-06-23
; PRIOR APPLICATION NUMBER: 60/213,634
; FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/214,926
; FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/261,320
; FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/264,353
; FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/273,879
; FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 09/887,271
; FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 6813
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6355
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-272A-6355

Query Match 35.4%; Score 41.8; DB 36; Length 2000;
Best Local Similarity 59.8%; Pred. No. 8.9;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 1 TAATACACACAAATTAATGTTGTTATAGCTTATATAATAATCTAGGGG 60
DB 633 TAATACAAACAAACATATTTTATTAATAGCTAATTTTCATGTT 574
QY 61 TTTAAGCGACACAAATTTATCGATAAATAATCTTTTAGACGCGACACAAATTTATA 117

Db 573 TTAGTGCCTGTTAAATTTTCTATATTAAATAGGAAATATCAAAATACTTATA 517

Search completed: February 28, 2004, 10:49:19
Job time : 1860.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:22:37 ; Search time 20.2561 Seconds
(without alignments)
981.354 Million cell updates/sec

Title: US-09-895-435A-2

Perfect score: 118

Sequence: 1 taatacaacaataataa.....agacgaacacaatttatag 118

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 164199 seqs, 84230614 residues

Total number of hits satisfying chosen parameters: 328398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents_NA_New:*
1: /cgn2_6/ptodata/1/pna/pct_NEW_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	31.4	26.6	545025	6	US-10-767-471-10717
C 2	31.4	26.6	1790242	6	US-10-767-471-10805
C 3	30.8	26.1	641	6	US-10-767-701-20322
C 4	30.6	25.9	201	6	US-10-767-471-38439
C 5	30.6	25.9	36106	6	US-10-767-471-10713
C 6	30.6	25.9	51995	6	US-10-767-471-10756
C 7	30.4	25.8	1014	5	US-08-830-230A-255
C 8	30.4	25.8	1047	5	US-08-830-230A-255
C 9	29.6	25.1	99445	6	US-10-767-471-10714
C 10	29.4	24.9	884	6	US-10-451-467A-695
C 11	29.2	24.7	201	6	US-10-767-471-13412
C 12	29.2	24.7	25346	6	US-10-767-471-10575
C 13	28.6	24.2	3872	5	US-08-673-389A-3
C 14	28.6	24.2	265113	6	US-10-767-471-10658
C 15	28.6	24.2	1126118	6	US-10-767-471-10615
C 16	28.4	24.1	201	6	US-10-767-471-27214
C 17	28.4	24.1	1758	6	US-10-417-884A-1840
C 18	28.4	24.1	99971	6	US-10-767-471-10561
C 19	28.4	24.1	144013	6	US-10-767-471-10668
C 20	28.4	24.1	380963	6	US-10-765-790-5
C 21	28.4	24.1	383432	6	US-10-765-790-34
C 22	28.2	23.9	170546	6	US-10-767-471-10703
C 23	28	23.7	500	1	PCT-US04-02000-576
C 24	28	23.7	231732	6	US-10-767-471-10869
C 25	28	23.7	234734	6	US-10-767-471-10827
C 26	28	23.7	435302	6	US-10-767-471-10552

C 27	27.8	23.6	201	6	US-10-767-471-23856
C 28	27.8	23.6	118356	6	US-10-767-471-10647
C 29	27.8	23.6	380963	6	US-10-765-790-5
C 30	27.8	23.6	546025	6	US-10-767-471-10717
C 31	27.6	23.4	201	6	US-10-767-471-42250
C 32	27.6	23.4	387	6	US-10-767-701-17975
C 33	27.6	23.4	500	1	PCT-US04-02000-300
C 34	27.6	23.4	1499	6	US-10-451-467A-79
C 35	27.6	23.4	76829	6	US-10-765-790-27
C 36	27.6	23.4	163548	6	US-10-767-471-10730
C 37	27.6	23.4	765267	6	US-10-767-471-10616
C 38	27.4	23.2	201	6	US-10-767-471-25904
C 39	27.4	23.2	1006	6	US-10-767-471-444
C 40	27.4	23.2	1157	6	US-10-767-471-443
C 41	27.4	23.2	1357	6	US-10-767-701-10249
C 42	27.4	23.2	12885	6	US-10-767-471-10657
C 43	27.4	23.2	47459	6	US-10-767-471-10701
C 44	27.4	23.2	166020	6	US-10-767-471-10806
C 45	27.4	23.2	265113	6	US-10-767-471-10658

ALIGNMENTS

RESULT 1
US-10-767-471-10717/c
; Sequence 10717, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001503
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10717
; LENGTH: 546025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(546025)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables
US-10-767-471-10717

Query Match 26.6%; Score 31.4; DB 6; Length 546025;
Best Local Similarity 57.7%; Pred. No. 12;
Matches 56; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 2 AATCAACACACATATTAATTGTTGTTAGTGTATTAATAATAATACTAGGGGT 61
Db 545424 AATCAAGCATACATTGTTGTTTATATTACCTGTTATCTATTAAATATTTGTT 545365
QY 62 TTAACGCAACACAAATTCGTAATAATAATACTTTT 98
Db 545364 TTATGCGCACCACTTAATCAATTCATTAAATGTTATT 545328
RESULT 2
US-10-767-471-10805/c
; Sequence 10805, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10805
; LENGTH: 1790242

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(1790242)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10805

Query Match      26.6%; Score 31.4; DB 6; Length 1790242;
Best Local Similarity 61.7%; Pred. No. 12;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 19 ATTCTGTTGTTAGTGTGTTATTAATAATAATAAATCTAGGGGTTTAAACGCAACACAATTT 78
Db 1587671 ATTTAAATGAATAATGATTCTAGTAAATAAGGAATAAATTGTATTAAGCAACAATATTT 1587612

Qy 79 ATCGATAAATAAATACCTTTTA 99
Db 1587611 ATGCACTRAAATAATATA 1587591

RESULT 3
US-10-767-701-20322
; Sequence 20322, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 20322
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB5048-003-R1-XP2-D10
US-10-767-701-20322

Query Match      26.1%; Score 30.8; DB 6; Length 641;
Best Local Similarity 55.7%; Pred. No. 14;
Matches 59; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 10 ACAATATAATTGTTGTTAGTGTGTTATTAATAATAATAAATCTAGGGGTTTAAACGCA 69
Db 175 AGAAGATTTTAAAGTTTATGTTACATTATATAAATAAATAAATAAATCTTGAAGTGAATACC 234

Qy 70 ACACAATTTATCGATAAATAAATACCTTTTAGACGCAACACAATTTA 115
Db 235 TCCAACTCTATGATATCTAGTTATTTAGAAAGAAATTAATTA 280

RESULT 4
US-10-767-471-38439
; Sequence 38439, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38439
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-38439
```

```
Query Match      25.9%; Score 30.6; DB 6; Length 201;
Best Local Similarity 62.3%; Pred. No. 15;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 17 TAATTGTTGTTAGTGTGTTATTAATAATAATAAATCTAGGGGTTTAAACGCAACACAAT 76
Db 116 TAATCGTGTGTTATATATCTTTAAATGGAATAAATAAATTTATTTATTTTAAAGCTATAAAA 175

Qy 77 TTATCGATAAATAAATA 93
Db 176 TTATAAATCAGTCATTA 192

RESULT 5
US-10-767-471-10713/c
; Sequence 10713, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10713
; LENGTH: 36106
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(36106)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables
US-10-767-471-10713

Query Match      25.9%; Score 30.6; DB 6; Length 36106;
Best Local Similarity 56.4%; Pred. No. 17;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 17 TAATTGTTGTTAGTGTGTTATTAATAATAATAAATCTAGGGGTTTAAACGCAACACAAT 76
Db 30733 TACITTTTGTGTTAAAGTATATATATGTTATATATATCTTTTGTATATAAAGTAT 30674

Qy 77 TTATCGATAAATAAATACTTTTAGACGCAACACAATTTATA 117
Db 30673 ATATATGTTATATATATATCTTTTGTATATAAAGTATA 30633

RESULT 6
US-10-767-471-10756
; Sequence 10756, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10756
; LENGTH: 51995
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10756

Query Match      25.9%; Score 30.6; DB 6; Length 51995;
Best Local Similarity 62.3%; Pred. No. 18;
Matches 48; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 17 TAATTGTTGTTAGTGTGTTATTAATAATAATAAATCTAGGGGTTTAAACGCAACACAAT 76
Db 39488 TAATCGTGTGTTATATCTTTAAATGGAATAAATAAATTTATTTATTTTAAAGCTATAAAA 39547
```



```
QY 77 TTATCGATAATAAATA 93
Db 39548 TTATAAATCAGTCATTA 39564

Query Match 25.8%; Score 30.4; DB 5; Length 1047;
Best Local Similarity 61.2%; Pred. No. 18;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

RESULT 7
US-09-830-230A-255/c
; Sequence 255, Application US/09830230A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR FILING DATE: 1998-06-18
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-09-03
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-255

Query Match 25.8%; Score 30.4; DB 5; Length 1014;
Best Local Similarity 61.2%; Pred. No. 17;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 26 TGTATTAGGTTTATAATAAATCTAGGGTTTAAAGCGCAACACAAATTTATCGATA 85
Db 254 TTTTAAAGCTGATAGCTATATAAAGCTAAAGGAATAGGCGATGATATATCGAAG 195

QY 86 AATAAATACCTTTAGACGCA 105
Db 194 AATAGATTGTTCTGACGTA 175

Query Match 25.8%; Score 30.4; DB 5; Length 1014;
Best Local Similarity 61.2%; Pred. No. 17;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

RESULT 8
US-09-830-230A-255/c
; Sequence 255, Application US/09830230A
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Lyme Disease Vaccines
; FILE REFERENCE: PB481US
; CURRENT APPLICATION NUMBER: US/09/830,230A
; CURRENT FILING DATE: 2001-09-27
; PRIOR FILING DATE: 1998-06-18
; PRIOR FILING DATE: 1997-09-03
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-07-22
; PRIOR FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 756
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 255
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-830-230A-255

Query Match 25.8%; Score 30.4; DB 5; Length 1047;
Best Local Similarity 61.2%; Pred. No. 18;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 26 TGTATTAGGTTTATAATAAATCTAGGGTTTAAAGCGCAACACAAATTTATCGATA 85
Db 254 TTTTAAAGCTGATAGCTATATAAAGCTAAAGGAATAGGCGATGATATATCGAAG 228

QY 86 AATAAATACCTTTAGACGCA 105
Db 227 AATAGATTGTTCTGACGTA 208

Query Match 25.8%; Score 30.4; DB 5; Length 1047;
Best Local Similarity 61.2%; Pred. No. 18;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

RESULT 9
US-10-767-471-10714/c
; Sequence 10714, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10714
; LENGTH: 99445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-10714

Query Match 25.1%; Score 29.6; DB 6; Length 99445;
Best Local Similarity 57.6%; Pred. No. 30;
Matches 53; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 26 TGTATTAGGTTTATAATAAATCTAGGGTTTAAAGCGCAACACAAATTTATCGATA 85
Db 25458 TCTTTTGAATTTACAAATAGAAAATAGAGATAAGATTAAACAAATTTAAAGCGGA 25399

QY 86 AATAAATACCTTTAGACGCAACACAAATTTATA 117
Db 25398 AATATATACCTTTAATAGAGCACAATAATCAAAA 25367

Query Match 24.9%; Score 29.4; DB 6; Length 884;
Best Local Similarity 54.1%; Pred. No. 30;

RESULT 10
US-10-451-467A-695/c
; Sequence 695, Application US/10451467A
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; FILE REFERENCE: YEA5T AND FUNGI
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-01-04
; PRIOR FILING DATE: 2001-01-09
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 695
; LENGTH: 884
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-695
```

Matches 60; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 1 TAATAACAACAATAATTTGGTTGTTATAGGTGTTATATAATAATAAATCTAGGG 60
Db 368 TAAATCTTCAGCATTTGTTTACTATGTTTATGATATATTCAGCTCTTTTGTG 309

QY 61 TTTAAGCGACAATAATTCGTAATAATAATACCTTTTAGACGCAACAA 111
Db 308 TTTATCATAATCTAGAAATCAATAGACAGATTATTCGGGATGGCACAA 258

RESULT 11
US-10-767-471-13412/c
; Sequence 13412, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13412
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-767-471-13412

Query Match 24.7%; Score 29.2; DB 6; Length 201;
Best Local Similarity 57.8%; Pred. No. 32;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 5 ACACACACATATTAATGTTGTTATAGGTGTTATATAATAATAATCTAGGGTTTA 64
Db 192 ACATGCCAGGTAACCTTACTTATTTTATAGAATTAAGACAAATTTAGGATTCA 133
QY 65 ACACACACAAATTTATCGATAATAATAATAC 94
Db 132 GCTGCACATAATTTGCAATTCATATTAC 103

RESULT 12
US-10-767-471-10575/c
; Sequence 10575, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10575
; LENGTH: 25346
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(25346)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10575

Query Match 24.7%; Score 29.2; DB 6; Length 25346;
Best Local Similarity 57.8%; Pred. No. 36;
Matches 52; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 5 ACACACAAATTAATGTTGTTATAGGTGTTATATAATAATAATCTAGGGTTTA 64
Db 12150 ACATGCCAGGTAACCTTACTTATTTTATAGAATAAGACAAATTTAGGATTCA 12091
QY 65 ACACACACAAATTTATCGATAATAATAATAC 94

Db 12090 GCTGCACATAATTTGTCGAATTCATATTAC 12061

RESULT 13
US-09-673-389A-3/c
; Sequence 3, Application US/09673389A
; GENERAL INFORMATION:
; APPLICANT: Jomaa, Hassan
; TITLE OF INVENTION: Process for identifying chemical active ingredients and
; TITLE OF INVENTION: active ingredients for inhibiting the
; TITLE OF INVENTION: 1-deoxy-D-xylulose-5-phosphate biosynthesis pathway
; FILE REFERENCE: 15514
; CURRENT APPLICATION NUMBER: US/09/673,389A
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: PCT/EP99/02463
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: DE19843279.8
; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: DE19816196.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: DE19828097.1
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: DE19825585.3
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: DE19831637.2
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: DE19831639.9
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: DE19831638.0
; PRIOR FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3872
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (126)..(3740)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(3872)
; FEATURE:
; NAME/KEY: mRNA
; LOCATION: (1)..(3872)
US-09-673-389A-3

Query Match 24.2%; Score 28.6; DB 5; Length 3872;
Best Local Similarity 59.0%; Pred. No. 47;
Matches 49; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 13 ATATTAAATGTTGTTATAGGTGTTATATAATAATAATCTAGGGTTTAACGCAACA 72
Db 3509 ATGTTTATTTGTTTATTATACATGATCTCATATTTTATCTAAAGGATTTAAAAATAT 3450
QY 73 CAATTATCGATAATAATAACT 95
Db 3449 CATATCAACAATTTGAAAARGAAT 3427

RESULT 14
US-10-767-471-10658
; Sequence 10658, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 10658
; LENGTH: 265113
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(265113)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10658

Query Match      24.2%; Score 28.6; DB 6; Length 265113;
Best Local Similarity 55.6%; Pred. No. 51;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      4  TACACACACATATTAATGTTGTTAGTGTATTAATAATAATAATAATACTAGGGGTTT 63
Db      230103 TACTATCAAAATAAATAATTATAGTACTCTCTCAATCCAGGAAATGTAAGGTTT 230162

QY      64  AACGCAACACAAATTCGATTAATAATACTTTTAGAC 102
Db      230163 AGCACTTCATATGTTTCATTACTAAATATTTTATAC 230201

RESULT 15
US-10-767-471-10615/c
; Sequence 10615, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10615
; LENGTH: 1126118
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1126118)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10615

Query Match      24.2%; Score 28.6; DB 6; Length 1126118;
Best Local Similarity 53.0%; Pred. No. 46;
Matches 61; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY      3  ATCAACACACATATTAATGTTGTTAGTGTATTAATAATAATAATACTAGGGGTT 62
Db      710818 ATAAACAGAAATTCGCAATGCTGGAGAAATAATACTACTTGAGAAATGAGAGCAA 710759

QY      63  TAACGCAACAAATTTATCGATTAATAATACTTTTAGCGCAACAAATTTATA 117
Db      710758 TTGACTATAGACTTCTTGCCCAATAATAATACTCTCGGAAGCACTACAAATAACA 710704

Search completed: February 28, 2004, 10:50:41
Job time : 29.2561 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 28, 2004, 07:05:57 ; Search time 1392.53 Seconds
(without alignments)
2530.460 Million cell updates/sec

Title: US-09-895-435A-2

Perfect score: 118

Sequence: 1 taatacaacaataattataa.....agacgaacacaattattag 118

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_esthum.*

4: em_esthum.*

5: em_esthum.*

6: em_esthum.*

7: em_esthum.*

8: em_esthum.*

9: gb_est1.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estfun.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_fun.*

22: em_gss_man.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41.6	35.3	420	12	BM891761
2	39.2	33.2	565	9	AV882598
3	39.2	33.2	1201	13	BX377054
4	39	33.1	713	28	BH986499

C 5	38.2	32.4	1201	13	BX361166
C 6	37.8	32.0	866	28	AQ258984
C 7	37.2	31.5	682	29	CG259325
C 8	37.2	31.5	901	29	CG456652
C 9	37.2	31.5	902	29	CG456652
C 10	37.2	31.5	926	29	CG37547
C 11	37.2	31.5	971	29	CG116679
C 12	37.2	31.4	531	12	BX369046
C 13	37.2	31.4	928	29	CNS00DKY
C 14	36.4	30.8	394	28	BH358603
C 15	36.4	30.8	433	28	AZ144898
C 16	36.4	30.8	639	14	CD718817
C 17	36.4	30.8	648	14	CD006182
C 18	36.4	30.8	677	28	AQ450336
C 19	36.4	30.5	837	29	CG356812
C 20	36.4	30.5	844	29	CG127867
C 21	36.4	30.5	987	29	CNS014PQ
C 22	35.8	30.3	487	14	CF606387
C 23	35.8	30.3	546	14	CD714159
C 24	35.8	30.3	580	10	BF043295
C 25	35.8	30.3	587	14	CD007820
C 26	35.8	30.3	599	14	CD007301
C 27	35.8	30.3	599	14	CD713999
C 28	35.8	30.3	602	14	CD004834
C 29	35.8	30.3	616	14	CD005930
C 30	35.8	30.3	638	14	CD008686
C 31	35.8	30.3	643	14	CF762291
C 32	35.8	30.3	646	14	CD008555
C 33	35.8	30.3	649	14	CD009563
C 34	35.8	30.3	649	14	CD009635
C 35	35.8	30.3	654	14	CD713887
C 36	35.8	30.3	656	14	CD008817
C 37	35.8	30.3	656	14	CD718517
C 38	35.8	30.3	658	14	CD006026
C 39	35.8	30.3	911	28	BH153166
C 40	35.6	30.2	831	29	CG204053
C 41	35.6	30.2	929	29	CG172162
C 42	35.6	30.2	957	29	CG462030
C 43	35.4	30.0	632	9	AI007994
C 44	35.4	30.0	682	28	AQ450315
C 45	35.4	30.0	776	10	BE111850

ALIGNMENTS

BM891761 420 bp mRNA linear EST 11-MAR-2002
sam43f11.y1 Gm-c1068 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1068-7581 5', mRNA sequence.

ACCESSION BM891761 GI:19346873
VERSION BM891761
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 420)
AUTHORS Shoemaker R., Keim P., Vodkin L., Erpelting J., Coryell V.,
Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J.,
Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M.,
Bowers Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N.,
Schurk R., Ritter E., Kohn S., Shin T., Jackson Y., Cardenas M.,
McCann R., Waterston R. and Wilson R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Db 1055 AYTAGACATCGMATTATTATWWWGTHANGKWITWWWAGMMCMCTTTT 1106

RESULT 4
BH986499
LOCUS
DEFINITION
oe110a12.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION
BH986499
VERSION
BH986499.1 GI:23520640
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 713)
Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE
Whole genome shotgun reads from Brassica oleracea
JOURNAL
Unpublished (2002)
COMMENT
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: oe110 row: h column: 12
Seq primer: -210ppor forward
Class: shotgun
High quality sequence start: 115
High quality sequence stop: 551.
FEATURES
Location/Qualifiers
1..713
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note=Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea F0100DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

Query Match 33.1%; Score 39; DB 28; Length 713;
Best Local Similarity 65.5%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 15 ATTAAATGTGTGTTAGTGGTTATATAATAATAATCTAGGGTTTAAACGCAACACA 74
Db 107 ATTAATAATATGTATTAAATTTTAAATAATAATACCAATTTAAAGTTTAAAAATATA 166
QY 75 ATTATCGATAATAATACTTTTAGA 101
Db 167 AATAATTATATATAAAAAATTATATA 193
RESULT 5
BX361166
LOCUS
DEFINITION
BX361166 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1081YM05 5-PRIME, mRNA sequence.
ACCESSION
BX361166
VERSION
BX361166.1 GI:30384495
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

Query Match 33.1%; Score 39; DB 28; Length 713;
Best Local Similarity 65.5%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 15 ATTAAATGTGTGTTAGTGGTTATATAATAATAATCTAGGGTTTAAACGCAACACA 74
Db 107 ATTAATAATATGTATTAAATTTTAAATAATAATACCAATTTAAAGTTTAAAAATATA 166
QY 75 ATTATCGATAATAATACTTTTAGA 101
Db 167 AATAATTATATATAAAAAATTATATA 193
RESULT 5
BX361166/c
LOCUS
DEFINITION
BX361166 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1081YM05 5-PRIME, mRNA sequence.
ACCESSION
BX361166
VERSION
BX361166.1 GI:30384495
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

TITLE
JOURNAL
COMMENT
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6196.r,
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CSOD1081AG03QPl.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1081YM05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 32.4%; Score 38.2; DB 13; Length 1201;
Best Local Similarity 43.6%; Pred. No. 38;
Matches 51; Conservative 23; Mismatches 43; Indels 0; Gaps 0;
QY 2 AATACACACACATATTATTTGTTTATTAGTGTGTATAATAATAATAATACTAGGGGT 61
Db 1195 AAAAAAATAAAAAAANKWATTTTGTGDKGKGTGKGTWAAAWAAAAATRGSGG 1136
QY 62 TTACGCAACACATTTATCGATAATAATAATACTTTTAGACGCAACACATTTATAG 118
Db 1135 AGKGGGGGTRWRAATGTATATAAAAAAAGGTGCGAARAKAATATGTTWAAAR 1079

ORIGIN
Query Match 32.4%; Score 38.2; DB 13; Length 1201;
Best Local Similarity 43.6%; Pred. No. 38;
Matches 51; Conservative 23; Mismatches 43; Indels 0; Gaps 0;
QY 2 AATACACACACATATTATTTGTTTATTAGTGTGTATAATAATAATAATACTAGGGGT 61
Db 1195 AAAAAAATAAAAAAANKWATTTTGTGDKGKGTGKGTWAAAWAAAAATRGSGG 1136
QY 62 TTACGCAACACATTTATCGATAATAATAATACTTTTAGACGCAACACATTTATAG 118
Db 1135 AGKGGGGGTRWRAATGTATATAAAAAAAGGTGCGAARAKAATATGTTWAAAR 1079
RESULT 6
LOCUS
DEFINITION
AQ258984 866 bp DNA linear GSS 23-OCT-1998
nbxb0021P09r CUGI Rice BAC Library Oryza sativa (japonica
cultivar-group) genomic clone nbxb0021P09r, genomic survey
sequence.
ACCESSION
AQ258984
VERSION
AQ258984.1 GI:3783466
KEYWORDS
GSS.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 866)
Wing,R.A. and Dear,R.A.
AUTHORS
A BAC End Sequencing Framework to Sequence the Rice Genome
TITLE
Unpublished (1998)
JOURNAL
COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGACACAGCTATGACCATG
Class: BAC ends
High quality sequence stop: 204.
Location/Qualifiers
1..866
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

FEATURES
source
1..866
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

```

/clone="nbxb0021P09r"
/tissue type="Leaf"
/lab host="E. coli DH10B"
/clone lib="CUGI Rice BAC Library"
/notes"Vector: pBelOBAC11; Site 1: HindIII; Site 2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

ORIGIN
Query Match 32.0%; Score 37.8; DB 28; Length 866;
Best Local Similarity 61.2%; Pred. No. 52; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 38;

QY 1 TAATACACACAAATATTGTTGTTAGGTGTTATAATAATAATACTAGGGG 60
    |||||
Db 752 TAATATAATATAATTATAATATTATTATTATAAATAATAATAATAATAATA 693
    |||||

QY 61. TTAAACGACACAAATTCGTAATAATAATACTTTT 98
    |||||
Db 692 ATTAATATATATATTTTAAATTAATAATAATAATT 655
    |||||

RESULT 7
LOCUS CG259325 682 bp DNA linear GSS 25-AUG-2003
DEFINITION OGWK85TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0610P01,
genomic survey sequence.
ACCESSION CG259325
VERSION CG259325.1 GI:34167746
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 682)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nurnberg, A., Robbins, D., and Lakey, N.
CONSORTIUM Consortium for Maize Genomics
TITLE Unpublished (2002)
JOURNAL Other GSSs: OGWK85TV
COMMENT Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..682
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0610P01"

FEATURES
source

```

```

/clone lib="ZM 0.7 1.5 KB"
/notes"Vector: pBCKK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 31.5%; Score 37.2; DB 29; Length 682;
Best Local Similarity 61.2%; Pred. No. 77; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 38;

QY 2 AATACACACAAATATTGTTGTTAGGTGTTATAATAATAATACTAGGGGT 61
    |||||
Db 34 AATATAATAATATTATTATTATTATGATTGTTTATCATTTAGTTTGTGT 93
    |||||

QY 62 TTAAACGACACAAATTCGTAATAATAATACTTTTA 99
    |||||
Db 94 TTAACCTTAAATTTTATATTTTAAATAATAATAATTGA 131
    |||||

RESULT 8
LOCUS CG456652/c 901 bp DNA linear GSS 17-SEP-2003
DEFINITION PUFLD31TDB ZM 0.6 1.0 KB Zea mays genomic clone ZMMBta0675E14,
genomic survey sequence.
ACCESSION CG456652
VERSION CG456652.1 GI:34841652
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 901)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and
Bennetzen, J.
CONSORTIUM Maize Genomics Consortium
TITLE Unpublished (2003)
JOURNAL Other GSSs: PUFLD31TB PUFLD31TD PUFLD31TBB
COMMENT Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..901
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBta0675E14"
/clone_lib="ZM 0.6 1.0 KB"
/notes"Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

FEATURES
source

```

```

LOCUS      CG456613                902 bp    DNA    linear    GSS 17-SEP-2003
DEFINITION PUF3D31TB.ZM.0.6.1.0.KB.Zea mays genomic clone ZMMBTA0675E14,
            genomic survey sequence.
ACCESSION  CG456613
VERSION    CG456613.1 GI:34841613
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 902)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUF3D31TB.PUF3D31TD.PUF3D31TDB
            Contact: Cathy Whitelaw
            TIGR Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES   Location/Qualifiers
            source          1..902
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone="ZMMBTA0675E14"
                        /clone_lib="ZM.0.6.1.0.KB"
                        /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                        Cot selected genomic DNA library"
ORIGIN
Query Match      31.5%; Score 37.2; DB 29; Length 902;
Best Local Similarity 61.2%; Pred. No. 70;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY  2 AATACACACAAATATTATGTTGTTGTTAGGTGTTATATAATAATAAATCTAGGGGT 61
Db  |||||
QY  246 AAAATAATAAATATTATTTATTTTATGATTTGTTTATCATTTAAGTTAGTTGTGT 305
Db  |||||
QY  62 TTAACGCAACAAATTTATCGATAAATAACTTTTA 99
Db  |||||
QY  306 TTAACCTAAATTTTATTTATTTTAAATAAATATTGGA 343
Db  |||||

RESULT 10
LOCUS      CG367547                926 bp    DNA    linear    GSS 26-AUG-2003
DEFINITION OGI:AH17TV.ZM.0.7.1.5.KB.Zea mays genomic clone ZMMBMA0716D10,
            genomic survey sequence.
ACCESSION  CG367547
VERSION    CG367547.1 GI:34284814
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 926)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
            Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
TITLE     Consortium for Maize Genomics
JOURNAL   Unpublished (2002)
COMMENT   Other GSSs: OGI:AH17TH
            Contact: Cathy Whitelaw
            TIGR

```

```

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES   Location/Qualifiers
            source          1..926
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone="ZMMBMA0716D10"
                        /clone_lib="ZM.0.7.1.5.KB"
                        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                        methylation filtered genomic DNA library"
ORIGIN
Query Match      31.5%; Score 37.2; DB 29; Length 926;
Best Local Similarity 61.2%; Pred. No. 69;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY  2 AATACACACAAATATTATGTTGTTGTTAGGTGTTTATATAATAATAAATCTAGGGGT 61
Db  |||||
QY  639 AAAATATAAATATTATTTATTTTATGATTTGTTTATCATTTAAGTTAGTTGTGT 698
Db  |||||
QY  62 TTAACGCAACAAATTTATCGATAAATAACTTTTA 99
Db  |||||
QY  699 TTAACCTAAATTTTATTTATTTTAAATAAATATTGGA 736
Db  |||||

RESULT 11
LOCUS      CG116679                971 bp    DNA    linear    GSS 20-AUG-2003
DEFINITION PUF3D31TB.ZM.0.6.1.0.KB.Zea mays genomic clone ZMMBTA0675E14,
            genomic survey sequence.
ACCESSION  CG116679
VERSION    CG116679.1 GI:34000116
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 971)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PUF3D31TD
            Contact: Cathy Whitelaw
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-5843
            Fax: 301-838-0208
            Email: whitelaw@tigr.org
            Seq primer: TR
            Class: sheared ends.
FEATURES   Location/Qualifiers
            source          1..971
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /strain="B73"
                        /db_xref="taxon:4577"
                        /clone="ZMMBTA0675E14"
                        /clone_lib="ZM.0.6.1.0.KB"
                        /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                        Cot selected genomic DNA library"
ORIGIN
Query Match      31.5%; Score 37.2; DB 29; Length 971;
Best Local Similarity 61.2%; Pred. No. 68;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```


availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.choi.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tftp/bac_ends/rat/bac_end_intro.html
Plate: 163 Row: N Column: 14
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers

1. 394
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clones="CH230-163N14"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN

Query Match 30.8%; Score 36.4; DB 28; Length 394;
Best Local Similarity 61.7%; Pred. No. 1.4e+02;
Matches 58; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 4 TACACACAAATTAATGTTGTTGTTATAGGTGTTATATAAATAAATCTAGGGGTTT 63
DB 232 TACATCTCCAACTTATTTGTTGTTTATTTTAAAGTTATATATAAGAGTG 173
QY 64 AACGCAACACAAATTTATCGATAAATAATCTTT 97
DB 172 TTCCAGGATCTTTCAATAGCAATTAATCTTT 139

RESULT 15

AZ144898

LOCUS

DEFINITION SP.0043 Al H01 T7 Strongylocentrotus purpuratus, purple sea urchin,
sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone Plate=43 Col=1 Row=0, genomic survey sequence.

ACCESSION

AZ144898

VERSION

GSS.

SOURCE

ORGANISM

Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE

AUTHORS

1 (bases 1 to 483)
Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T.,
Wray,G.A., Etesohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H.
and Hood,L.

A sea urchin genome project: Sequence scan, virtual map, and
additional resources

JOURNAL

MEDLINE

PUBMED

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)
10920195
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 43 row: O column: 1
Seq primer: T7

Class: BAC ends

High quality sequence stop: 483.

FEATURES

source

Location/Qualifiers
1. 483

/organism="Strongylocentrotus purpuratus"

/mol_type="genomic DNA"
/db_xref="taxon:7668"
/clone="Plate=43 Col=1 Row=0"
/clone_lib="Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BAC3.6; BAC Clones in E-Coli
DH10B"

ORIGIN

Query Match 30.8%; Score 36.4; DB 28; Length 483;
Best Local Similarity 61.7%; Pred. No. 1.3e+02;
Matches 58; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 1 TAATACACACAATTAATTTGTTGTTATAGGTGTTATATAAATAAATCTAGGGG 60
DB 330 TAATATACAAATCTTAATTTTATGAAAAATGTTATAAAAAAATAAAGCGCAAGCAT 389
QY 61 TTTAAGCGCAACAATTTATCGATAAATAATAC 94
DB 390 GTGTCGGGACTTTTTTTTATTGATAAATTAATGC 423

Search completed: February 28, 2004, 09:05:52

Job time : 1396.53 secs